TECHNISCHE
UNIVERSITÄT
DARMSTADT

Security in Infrastructure-less and Decentralized Communication Networks

Location-based Intrusion Response and User-based Cooperative Decisions

Vom Fachbereich
Elektrotechnik und Informationstechnik
der Technischen Universität Darmstadt
zur Erlangung des Grades eines
Doktor-Ingenieurs (Dr.-Ing.)
genehmigte

Dissertationsschrift

von

Dipl.-Inform. André König
Geboren am 5. Februar 1978 in Darmstadt

Erstreferent: Prof. Dr.-Ing. Ralf Steinmetz
Korreferent: Prof. Klara Nahrstedt, Ph.D.
Korreferent: Prof. Dr.-Ing. Matthias Hollick
Vorsitz: Prof. Dr.-Ing. Hans Eveking

Tag der Einreichung: 19. Oktober 2010
Tag der Disputation: 20. Dezember 2010

Darmstadt, 2011
Hochschulkennziffer D17

Security in Infrastructure-less and Decentralized Communication Networks

Location-based Intrusion Response and User-based Cooperative Decisions

Zur Erlangung des Grades eines Doktor-Ingenieurs (Dr.-Ing.)
genehmigte Dissertation von Dipl.-Inform. André König, geboren am 5. Februar 1978 in Darmstadt
2011 – Darmstadt – D17

TECHNISCHE UNIVERSITÄT DARMSTADT

Fachbereich Elektrotechnik
und Informationstechnik

Fachgebiet Multimedia Kommunikation
Prof. Dr.-Ing. Ralf Steinmetz

Security in Infrastructure-less and Decentralized Communication Networks
Location-based Intrusion Response and User-based Cooperative Decisions
genehmigte Dissertation von Dipl.-Inform. André König, geboren am 5. Februar 1978 in Darmstadt

Tag der Einreichung: 19. Oktober 2010
Tag der Disputation: 20. Dezember 2010

Erstreferent: Prof. Dr.-Ing. Ralf Steinmetz
Korreferent: Prof. Klara Nahrstedt, Ph.D.
Korreferent: Prof. Dr.-Ing. Matthias Hollick
Vorsitz: Prof. Dr.-Ing. Hans Eveking

Technische Universitï£¡t Darmstadt
Fachbereich Elektrotechnik und Informationstechnik

Fachgebiet Multimedia Kommunikation (KOM)
Prof. Dr.-Ing. Ralf Steinmetz

Zusammenfassung

Infrastrukturlose Kommunikationssubstrate wie drahtlose multi-hop Ad-hoc-Netze in Kombination mit dezentralen, Peer-to-Peer-basierten Anwendungen ermöglichen durch ihre selbstorganisierende Natur und den Verzicht auf zentrale Instanzen den spontanen Aufbau digitaler Kommunikationsnetze. Mit Kommunikationsendgeräten, die gleichzeitig als Vermittlungsstationen dienen und so unabhängig von einer vorhandenen Kommunikationsinfrastruktur sind, umfassen mögliche Einsatzszenarien zum Beispiel den Aufbau von Kommunikationsnetzen in großräumigen Katastrophenszenarien. Aus dem Blickwinkel der Sicherheit bedeutet die Möglichkeit, selbstorganisierende Netze spontan aufzubauen, allerdings die Abhängigkeit von der Kooperation von Geräten aus mehreren Verwaltungsstrukturen ohne zentrale Kontrolle. Sicherheitsregelwerke, die zum Beispiel Zugriffe auf vertrauliche Dienste steuern, können während einer spontanen Zusammenarbeit nicht als vorhanden vorausgesetzt werden.

Diese Arbeit adressiert (1) die Aufrechterhaltung der Funktionalität des Netzes in Gegenwart von Geräten, die die kooperative Natur von infrastrukturlosen Netzen für Angriffe auf deren Verfügbarkeit ausnutzen und (2) das Erreichen von Schutzzielen wie Authentifizierung und Zugangskontrolle in Abwesenheit zentraler (vertrauter) Instanzen und vordefinierter Sicherheitsregelwerke.

Zu (1) wird ein neuartiger, positionsbasierter Verteidigungsmechanismus für infrastrukturlose Netze vorgestellt. Da Geräte hier keiner zentralen Kontrolle unterliegen, können Netzadressen mit geringem Aufwand geändert und adressbasierte Schutzmechanismen umgangen werden. Der positionsbasierte Ansatz, der in dieser Arbeit entwickelt wird, benutzt die physikalische Position statt der Netzadresse von Geräten zu deren Identifikation. Fehlverhaltende Geräte werden durch den Aufbau von Quarantänegebieten, die von der Kommunikation ausgenommen werden, aus dem Netz ausgeschlossen. Auf Basis analytischer Modelle und durch Simulationsstudien gewonnene Ergebnisse zeigen, dass auf diese Weise die Unanfälligkeit des Verteidigungsmechanismus gegenüber Änderungen der Netzadresse fehlverhaltender Geräte erreicht werden kann. Ein Nachteil des positionsbasierten Ansatzes ist der zur Aufrechterhaltung der Gesamtfunktionalität des Netzes notwendige Ausschluss gutartiger Geräte, die sich in physikalischer Nähe zu fehlverhaltenden Geräten befinden, aus dem Netz. Um diesem Effekt entgegenzuwirken, werden zwei Ansätze vorgestellt. Durch eine adaptive Sendeleistung der gutartigen Geräte wird die Ausdehnung der Quarantänegebiete reduziert; durch den Einsatz verzögerungstoleranter Anwendungen wird die (zeitverzögerte) Kommunikation von Geräten, die sich in Quarantänegebieten befinden ermöglicht. Auf Basis von Simulationsstudien gewonnene Ergebnisse zeigen, dass eine adaptive Sendeleistung zur Verbesserung des positionsbasierten Verteidigungsmechanismus in Szenarien mit geringer Mobilität genutzt werden kann. Durch den Einsatz verzögerungstoleranter Anwendungen kann der positionsbasierte Ansatz effektiv unterstützt werden, wenn Verzögerungen in Kauf genommen werden können.

Zu (2) werden in dieser Arbeit nutzerbasierte, kooperative Entscheidungen als Ersatz für zentrale, vertraute Instanzen und Sicherheitsregelwerke vorgestellt. Durch einen kooperativen Entscheidungsprozess basierend auf Schwellwertkryptographie wird verhindert, dass kryptographische Operationen wie das Signieren von Zertifikaten, die Zugriff auf vertrauliche Dienste gewähren, von einem einzelnen, möglicherweise kompromittierten Gerät ausgeführt werden können. Durch das direkte Einbeziehen von Nutzern in spontane Entscheidungsprozesse werden sicherheitsrelevante Entscheidungen ohne vordefinierte Sicherheitsregelwerke ermöglicht. Um die Anzahl der in einen Entscheidungsprozess involvierten Nutzer und die Häufigkeit, mit der ein bestimmter Nutzer beteiligt ist, zu minimieren, werden verschiedene Interaktionsschemata von Anfragendem und (potentiellen) Entscheidungsträgern verglichen. Zur Minimierung der in eine Entscheidung involvierten Nutzer werden analytische Modelle erstellt, die es erlauben, den Entscheidungsprozess zu steuern. Auf Basis einer prototypischen Implementierung in zwei Experimentalplattformen gewonnene Ergebnisse zeigen die Anwendbarkeit der Interaktionsschemata und die Korrektheit der analytischen Modelle.

Abstract

Infrastructure-less communication substrates like multi-hop wireless mobile ad hoc networks in combination with applications based on decentralized networks like peer-to-peer networks facilitate establishing digital communication services in a spontaneous way. Envisioned application scenarios include, for example, enabling communication in large-scale disaster scenarios. Here, a preexisting communication infrastructure might not be available. Thus, devices have to act both as communication endpoints and routers. Built upon the paradigm of self-organization, the functionality of both infrastructure-less and decentralized networks is based on the cooperation of the devices forming the network and on the abandonment of fixed, central instances. From the perspective of security, being able to establish self-organizing networks in a spontaneous way is, however, paid for by being dependent on the cooperation of devices from many administrative domains that are beyond a central control. Further, the availability of security policies controlling, for example, access to restricted resources can not be assumed during spontaneous interactions.

In this thesis, we address two major resulting challenges of (1) maintaining the functionality of the network in presence of devices that exploit the cooperative nature of infrastructure-less networks to launch attacks on network availability and (2) achieving security objectives like authentication and access control in the absence of a central (trusted) instance and predefined security policies.

Regarding Challenge (1), we present a novel, location-based intrusion response mechanism for infrastructure-less networks. Since devices in infrastructure-less networks are beyond a central control, changing network addresses of devices and, thus, circumventing conventional address-based intrusion response solutions is possible with little effort. The location-based intrusion response approach we develop within this thesis, instead, uses the physical location of devices as an identifier. Misbehaving devices are excluded from the network by establishing quarantined areas void of communication at locations where misbehavior is detected. Our results based on analytical modeling and simulation studies show that, this way, we render the intrusion response mechanism insusceptible to changes in addresses of misbehaving nodes. On the downside, benign devices located in close physical proximity to misbehaving nodes are, for the sake of overall network survivability, excluded from the network along with misbehaving nodes. To mitigate this effect, we propose two approaches based on (a) adaptive transmission power of devices to minimize the size of quarantined areas and (b) harnessing delay tolerance of applications to enable (delayed) communication of benign devices located within quarantined areas. Our results based on simulation studies show that an adaptive transmission power can improve the location-based intrusion response approach in scenarios with low node mobility. By harnessing delay tolerance, we are able to effectively support the location-based intrusion response at the cost of increased transmission delays.

Regarding Challenge (2), we present user-based, cooperative decisions as a replacement for central (trusted) instances and security policies. By introducing a cooperative decision process based on threshold cryptography, we prevent that cryptographic operations like signing certificates that grant access to restricted resources can be performed by a single, possibly compromised device. By involving (authorized) users directly in the decision process, we enable decisions on security-related requests during spontaneous interactions without predefined security policies. When involving users directly in decision processes, obviously, the number of users, as well as the frequency at which one particular user is requested, have to be minimized. To achieve these requirements, we discuss different interaction schemes between a user issuing a security-related request and the (potential) users taking part in the decision process. Subsequently, we present analytical models serving as tools for governing the decision process, in order to minimize the number of users involved in a decision. Our results obtained from a prototype for user-based, cooperative decisions deployed in two testbeds show the applicability of the interaction schemes as well as the correctness of the analytical models.

Contents

1 Introduction

A legitimate start for a PhD thesis on communication networks would be something like 'In the recent years, digital communication networks have found their way into many areas of our daily lives.'. However, for two reasons, the communication networks that are in the focus of this thesis render this start inappropriate.

First, considering the 'way', in terms of communication networks, we usually refer to communication infrastructures such as digital subscriber lines (DSL) or fiber-based broadband last mile connections providing triple-play services to private households. On national scale, we see towards all-IP tending backbone networks of communication service providers. On global scale, we have satellites and submarine cables of international carriers. Internet exchange points like Amsterdam Internet Exchange (AMS-IX) built the bridges between national and international providers. Yet, based on pure wireless communication, we are able to establish communication substrates that live without any of these infrastructure components. When doing so, we also have to rethink the way applications for these infrastructure-less environments are designed. Due to the potential mobility of nodes, the lack of a fixed infrastructure, and the wireless data transmission, the topology of infrastructure-less communication substrates is changing constantly. Hence, single nodes or groups of nodes may experience disconnections frequently. Deploying a centralized application on top of an infrastructure-less communication substrate could, thus, lead to long periods of unavailability of the central instance and the services offered. Applications based on decentralized networks, instead, are able to deal with the challenging conditions of an infrastructure-less communication substrate. Operating without central instances and, like infrastructure-less networks, built upon the paradigm of self-organization, decentralized networks can adapt to a constantly changing network topology and maintain the availability of services. From the technical perspective, regarding the 5-layer model of communication networks presented in [104], we consider infrastructure-less networks on network layer and below and decentralized applications on application layer, as shown in Figure 1.1. Altogether, being independent from preexisting infrastructures means a gain in flexibility for opening up new application areas for digital communication networks. Yet, especially in terms of security, it also means a loss of well-defined network borders and, thus, it means being confronted with new challenges that cannot be overcome by security mechanisms designed for infrastructure-based networks.

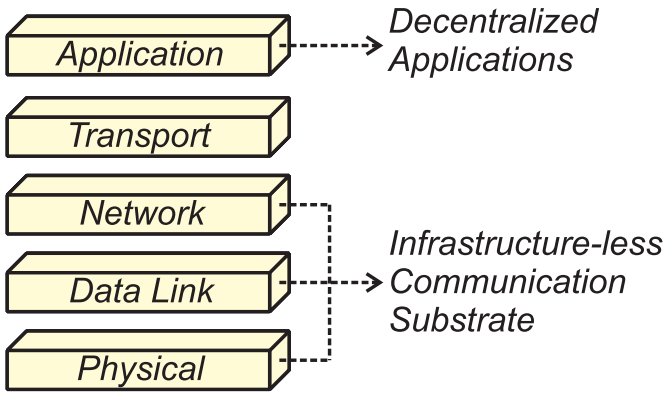


Figure 1.1: 5-layer model of communication networks showing where an infrastructure-less and decentralized nature is considered within this thesis

Second, the infrastructure-less and decentralized networks that are in the focus of this thesis, are not intended for deployment in our 'daily lives'. Due to the infrastructure-less nature, they cannot keep up with the performance of infrastructure-based networks that we demand for the communication

services we use. Network quality of service characteristics such as throughput, delay, and loss are clearly inferior. However, communication is of vital importance also in situations beyond our daily lives, in situations such as (large-scale) disaster scenarios, where communication infrastructures are not, or not anymore, available. Here, we can accept a degraded network performance as a trade-off for being able to communicate at all. Events like 9/11, the Elbe flood in 2002, or 2005 Hurricane Katrina showed that communication in large-scale disaster scenarios is still flawed. Regarding communication solutions for first responders, this is documented in [52, 101, 113, 64]. Research projects that consider this application scenario are, for example, DUMBO [50] and HiMONN [44]. Both projects focus on how mobile ad hoc networks can be deployed to support communication of first responders. However, besides supporting communication of first responders, keeping communication channels of affected civilians alive should not be neglected and, for example, can help to prevent mass hysteria.

In this thesis, we, thus, consider an infrastructure-less communication network as it might be established to enable communication in large-scale disaster scenarios for anyone affected. We assume that any device, of affected civilians as well as of emergency response units, that has wireless communication capabilities is used to form a scenario-wide network. Due to the resulting inherent vulnerability, effective security mechanisms are a basic requirement. In the following, we briefly identify properties of infrastructure-less and decentralized communication networks. We shortly summarize the state of the art and derive implications of the infrastructure-less and decentralized nature on security.

1.1 Infrastructure-less Networks and Security

The basic functionality of an infrastructure-less network is based on the cooperation of the devices forming the network. To enable wireless communication of devices that are not in direct transmission range, intermediate nodes act as routers and relay the communication appropriately. In this thesis, we use multi-hop wireless mobile ad hoc networks based on the IEEE 802.11 standard as particular instantiation of infrastructure-less networks. Yet, the security mechanisms developed can be adopted to other kinds of infrastructure-less networks like IEEE 802.16-based mesh networks, sensor networks, or vehicular ad hoc networks.

1.1.1 Implications of an Infrastructure-less Nature on Security

An infrastructure-less nature as described above brings along manifold implications on network security. Attacks enabled by the infrastructure-less nature are possible on each layer of the 5-layer model shown in Figure 1.1, in addition to well-known attacks on infrastructure-based communication systems. Contemporary security mechanisms like gateways or firewalls cannot be transferred (without adaptations) from infrastructure-based networks due to the lack of well-defined network borders that could be protected by theses mechanisms.

Attacks on the network layer of mobile ad hoc networks are, in general, based on injecting false routing and/or forwarding information into the network. One particular example of an attack on network layer is the black hole attack. Comparable to a black hole in terms of astronomy, a black hole in a mobile ad hoc network attracts and drops network traffic instead of forwarding it to the receiver it is intended for. The attracting effect is achieved by injecting false routing information that, in terms of the routing metrics used by the particular routing protocol, makes routes via the black hole appear more attractive than routes via benign nodes. Due to the severe and well quantifiable effects on network performance, we use the black hole attack to exemplify misbehavior in mobile ad hoc networks in this thesis. However, the applicability of the security mechanisms developed in this thesis is neither limited to the black hole attack in particular, nor to attacks on network layer in general.

1.1.2 State of the Art in Security for Infrastructure-less Networks

Security mechanisms (for mobile ad hoc networks) can be categorized into preventive and reactive mechanisms. Preventive mechanisms, which are not in the focus of this thesis, are, in general, based on means of cryptography to prevent false routing/forwarding information from being injected into the network. Although designed carefully, successful attacks on most preventive security mechanisms are discovered after a short lifetime.

In case preventive security mechanisms fail, reactive measures must be taken as a second line of defense. For our needs, reactive measures consist of intrusions detection systems combined with intrusion response mechanisms. In ad hoc networks, intrusion detection on network layer is commonly based on devices operating in promiscuous mode, monitoring the routing/forwarding behavior of neighbor nodes in transmission range. This approach was first presented in [65]. Considering the example of a black hole attack and assuming bidirectional wireless links, a node is able to check whether packets sent to a neighbor node for further forwarding are relayed correctly. Based on this information, nodes can be rated as benign or misbehaving. Subsequently, intrusion response mechanisms can be used to exclude misbehaving nodes from the network. Several schemes for intrusion detection and intrusion response in mobile ad hoc networks were proposed by now. Of relevance for our work is that mostly, like in [7, 13, 67], these schemes work in an address-based way. Misbehaving nodes are identified and excluded from the network based on their network addresses. However, evading address-based solutions is possible with little effort in a spontaneously established mobile ad hoc network consisting of devices from different administrative domains that, during operation, are beyond a central control. Also, a misbehaving node does not necessarily have to be a logical part of the network, that is, having a network address is not a requirement for attacking the network. Jamming attacks, as an example, can be performed without having joined the network logically and, thus, can not be thwarted by address-based intrusion response approaches.

1.2 Decentralized Networks and Security

In client/server systems, Internet browsers and web servers being a prominent example, providers and consumers of services are strictly separated. Services are provided by central servers and consumed by clients. In decentralized systems, in contrast, the strict separation of providers and consumers is broken up. In this thesis, we use peer-to-peer systems as particular instantiation of decentralized systems. Here, services are offered and consumed in an equal way by the nodes forming the network - the peers.

1.2.1 Implications of a Decentralized Nature on Security

From the perspective of security, waiving central instances means waiving trusted servers on which we rely for security services such as authentication and access control in centralized networks. Further, when peer-to-peer systems are established in a spontaneous way, for example, to enable communication in large-scale disaster scenarios, the availability of security policies defining access rights for restricted services cannot be assumed.

In general, traditional means for offering authentication and access control like Kerberos [74] are based on trusted servers and databases containing predefined access rights. As an example, we consider a scenario with a server offering restricted services, a client requesting access to a service, and a trusted server responsible for deciding on access requests. In this case, the client that requests access to a restricted service issues this request to the trusted server. If the database storing the access rights contains a corresponding entry, the trusted server replies with a cryptographipcally signed certificate that grants access to the requested service. The client then presents this certificate to the server offering the requested service that, after verifying the authenticity of the certificate, permits access. In peer-to-peer

systems, without central, trusted instances and security policies, security services cannot be offered this way. Thus, replacements for both central, trusted instances and security policies are required.

1.2.2 State of the Art in Security for Decentralized Networks

To offer security services in peer-to-peer systems, threshold cryptography can be used as replacement for central trusted instances. Here, a cryptographic key is distributed among multiple peers. This is achieved by choosing a polynomial such that it contains the key to be shared. The key shares then correspond to the polynomial at arbitrary index values (except that of the key itself). This approach was first presented in [95]. With the key shares, peers are able to produce partial signatures for certificates that, for example, grant access to restricted resources. By Lagrange interpolation, a full, valid signature can be computed only if a sufficient number of partial signatures is combined. Thus, no single, possibly compromised peer is able to decide on security-related requests.

The application of threshold cryptography in peer-to-peer systems (and mobile ad hoc networks) is well investigated. The foundation was laid in [73], where different threshold signature schemes are compared with respect to their performance in controlling access to closed user groups within P2P Systems. Performance is measured in terms of basic operation costs, that is, the time needed to produce partial signatures, and join time, that is, the amount of time a new peer needs to join a closed user group. In general, the focus of related work is set on the evaluation of the performance of different threshold cryptography schemes. Subject of the evaluation are mostly technical metrics such as computing time required, energy consumed, or data overhead generated. To the best of our knowledge, no attention was paid so far on how a lack of security policies and the resulting challenges can be overcome.

1.3 Contributions

In this thesis, we address the challenges in offering security in mobile ad hoc networks and peer-to-peer system mentioned above. We present (1) an intrusion response mechanism for mobile ad hoc networks that uses the physical location instead of network addresses to identify misbehaving devices and (2) user-based cooperative decisions in peer-to-peer networks as replacement for central, trusted instances and security policies.

1.3.1 Location-based Intrusion Response in Infrastructure-less Networks

Using the physical location of devices instead of network addresses, the location-based intrusion response mechanism we present in this thesis excludes misbehaving devices from the network by establishing quarantined areas at locations where misbehavior is detected. Benign devices being aware of their quarantined location stop communicating as long as they are quarantined. This way, we make quarantined areas void of communication, thus preventing communication from entering or leaving quarantined areas. As a consequence, existing routes that cross boundaries of quarantined areas become invalid and, since also control traffic responsible for establishing new routes neither enters nor leaves quarantined areas, newly established routes avoid quarantined areas.

For obvious reasons, quarantined areas have to cover at least the transmission/reception range of devices. Further, the location-based intrusion response is applicable only in scenarios as outlined above, consisting of a large number of devices, since benign devices located in physical proximity to misbehaving devices are excluded from the network for the sake of overall network survivability. Existing testbeds are of a small-scale compared to both the transmission/reception range of devices and the number of devices they consist of [53]. A testbed-based evaluation of the location-based intrusion response is, thus, hardly realizable. To evaluate the performance of the location-based intrusion response we, therefore, combine simulation studies and analytical modeling. This way, we aim at achieving a best

possible coverage of the parameter space and at proving the results obtained in simulation studies and by analytical models mutually plausible. Directly comparing an address-based solution with the location-based intrusion response at the example of a black hole attack, our results show that the location-based approach is insusceptible to changes in addresses of misbehaving nodes. To evade the location-based intrusion approach, misbehaving devices have to leave quarantined areas which is physically bounded by the speed of devices. Although we use the black hole attack to evaluate the performance of the location-based intrusion response, the applicability is not limited to this particular attack, nor to attacks on network layer in general. Also attacks on other layers, like tampering with medium access or transport protocols, once detected and localized, can be thwarted by the location-based intrusion response.

To mitigate the drawback of benign devices being excluded from the network along with misbehaving devices in close physical proximity, we present two approaches.

First, we apply an adaptive transmission power of devices to reduce the transmission/reception range, thus minimizing the required size for quarantined areas. Our results based on simulation studies show that this approach can be used in scenarios with low node mobility to improve the location-based intrusion response in terms of benign nodes affected. While in scenarios with a high node mobility, reducing transmission power still can be used to minimize the number of benign devices affected, the reduced transmission range causes frequent route breaks. The resulting effects counteract the gain of minimized quarantined areas and cause a degraded network performance.

As second approach to improve the location-based intrusion response, we harness delay tolerance of applications to enable delayed communication of quarantined devices. Results obtained in simulation studies show that we are able to nearly fully recover network functionality in terms of packet loss if we can accept increased transmission delays.

1.3.2 User-based Cooperative Decisions for Decentralized Networks

To provide security services like authentication and access control in peer-to-peer systems without a central, trusted instance and security policies, we present tools and interaction schemes for user-based cooperative decisions. To enforce the cooperation of peers for deciding on security-related requests, we distribute a cryptographic key by means of threshold cryptography. By involving users directly in a decision process, we are able to decide on security-related requests in the absence of security policies.

We first review existing threshold cryptography approaches subject to their applicability in our scenario, where we involve users in the decision process. Many threshold cryptography approaches require that the co-signers are known to each instance that contributes a partial signature before partial signatures can be generated. A peer that is part of a decision process and does not provide its partial signature then causes the overall decision process to fail. Thus, the unpredictable behavior of users, that is, the fact that users might not be able to provide a decision in a reasonable time, makes these threshold cryptography approaches unsuitable. We, therefore, require a threshold cryptography scheme that is able to handle failures of individual peers that are part of a decision process.

When involving users directly in decision processes, we have to keep in mind that the main interests of users in a disaster scenario are different ones. Thus, the frequency with which one particular user is involved as well as the overall number of users involved per decision have to be minimized. To achieve the first, we present different models for the interaction of a peer that issues a security-related request with the peers that are involved in the decision process. To achieve the latter, we present stochastic models describing the different interaction schemes, thus serving as tools for minimizing the number of users involved per decision during system operation. For model validation, we deploy a prototype for the user-based cooperative decisions in the PlanetLab [80] and the G-Lab [19] testbeds. The results obtained show the applicability of the interaction schemes and the correctness of the analytical models.

2 Basics and Related Work on Security in Mobile Ad Hoc Networks

In this chapter, we briefly review foundations of mobile ad hoc networks and basic as well as recent related work that has motivated our research. For a more comprehensive introduction, we refer to [31]. We focus on research related to the basic approach for location-based intrusion response. We present work on attacks and security mechanisms, location-based routing mechanisms and analytical models for mobile ad hoc networks. We further review work on adaptive transmission power and delay tolerant communication in mobile ad hoc networks which are the basis for the improvements of the location-based intrusion response. For a broader overview of security-related challenges, we refer to [15].

2.1 General Network Model

We consider a mobile ad hoc network that is established in a large-scale disaster scenario. Without loss of generality, we focus on mobile ad hoc networks established based on the IEEE 802.11 standard.

We assume that any device in the affected area featuring wireless communication capabilities is used to form the network. This includes devices of affected civilians as well as devices of on-site first responders and reconstruction units. We assume a resulting network size in the order of 10^3 nodes. While the location-based security mechanism developed in the following chapters can be applied to larger scenarios, it cannot be applied in small-scale scenarios, where routes consist of one or two hops only.

Sophisticated models for mobility patterns of first responders are introduced, for example, in [42, 93]. However, we assume that by establishing a scenario-wide network that consists of devices from different user groups, the resulting mobility pattern is a random movement. Regarding the security mechanisms developed, a random mobility can further be assumed as the worst-case.

Without loss of generality, we assume that the benign or malicious behavior of nodes is static. That is, nodes do not switch between benign and malicious behavior. We further assume that misbehaving nodes do not take part in the actual communication. A misbehaving node is neither source nor intended destination of a communication.

Regarding the communication pattern, we focus on unicast communication. However, the security mechanisms developed can be adapted to multicast and to broadcast communication. We assume that communication is sparse with respect to the number of nodes, the network consists of.

2.2 Foundations

In the following we present the details of the physical, data link, and network layer of mobile ad hoc networks required for the remainder of this thesis.

2.2.1 Physical Layer

Details of the physical layer that are of relevance for this thesis are antenna characteristics and models for wireless signal propagation. For a comprehensive overview, we refer to [103].

To propagate a signal wirelessly, energy has to be transmitted from a sender S to a receiver R. This is done by electromagnetic waves of a certain power P emitted and received by antennas.

In the simplest case, for an isotropic antenna, the power P_S is emitted spherically by the sending antenna and is, thus, distributed uniformly on the surface $A_{sphere} = 4\pi r^2$ of the sphere with radius r. This results in a power density

$$S_{iso} = \frac{P_S}{4\pi r^2}$$

at a distance r from the sending antenna. The power P_R received at the receiving antenna now depends on the effective area of the antenna

$$A_{eff,iso} = \frac{\lambda^2}{4\pi}$$

where λ denotes the wavelength of the signal. P_R now can be calculated as

$$P_{R,free,iso} = S_{iso} \cdot A_{eff,iso} = \frac{P_s}{4\pi r^2} \cdot \frac{\lambda^2}{4\pi}$$

which is a free space model that holds if the main lobe of the antenna is free of obstacles and if the distance between sender and receiver is small.

Since practical antennas do not emit spherically, but show planar peaks, we obtain an antenna gain G_S at the sender and G_R at the receiver for this omnidirectional antenna compared to the isotropic antenna. The resulting power density at a distance r from the sending antenna is then defined as

$$S_{omni} = \frac{P_S \cdot G_S}{4\pi r^2}$$

The effective area changes to

$$A_{eff,omni} = \frac{\lambda^2 \cdot G_R}{4\pi}$$

For an omnidirectional antenna, we, thus, obtain

$$P_{R,free,omni} = S_{omni} \cdot A_{eff,omni} = G_S \cdot G_R \cdot \frac{P_s}{4\pi r^2} \cdot \frac{\lambda^2}{4\pi}$$

which also is a free space model that holds if the main lobe of the antenna is free of obstacles and if the distance between sender and receiver is small.

We require the models to calculate a reduced transmission power of devices in order to reduce the size of quarantined areas while keeping the security level as high as possible. Reducing the transmitted power and, thus, the received power more than required has no negative effect on security. To keep the complexity of the evaluation on a feasible level for the simulation tool we, therefore, neglect effects of obstacles that mostly cause an additional signal attenuation.

If the distance between sender and receiver is large, effects of signals reflected from the ground causing additional attenuation but also amplification have to be considered. The corresponding two ray ground reflection model describes the power received at the receiver as

$$P_{R,tworay} = G_S G_R \frac{P_S h_S^2 h_R^2}{r^4}$$

where h_S and h_R denote the height of sender and receiver above the ground.

By setting $P_{R,free} = P_{R,tworay}$ and resolving to the distance r between sender and receiver, we obtain the cross-over distance

$$r_C = \frac{4\pi h_S h_R}{\lambda}$$

If the distance r between sender and receiver is below r_C, effects of ground reflection can be neglected and the power received at the receiver can be calculated with the free space model.

In our scenario, we assume a communication based on IEEE 802.11 wireless networks operating at $2.4GHz$ resulting in $\lambda = 0.125m$. Further, we assume a height of sender and receiver above ground that corresponds to devices held at human head level, thus $h_S \approx h_R \approx 1.6m$. The resulting cross-over distance is $r_C \approx 257m$, which is larger than the transmission range of $250m$ that we assume. For our calculations in Chapter 4, we, thus, apply the free space model.

2.2.2 Data Link Layer

A core challenge of wireless data transmission is to coordinate access to the wireless medium. In wired (local area) networks, all devices attached to the wire are able to hear whether an arbitrary device uses the wired medium. Protocols coordinating access to the medium can be based on a simple check whether the medium is in use before a transmission is started in combination with a collision detection.

In multi-hop mobile wireless ad hoc networks, this access control scheme is not applicable. If we consider the topology shown in Figure 2.1, then, due to the limited wireless transmission range, Device C may sense the medium as free, while Device A uses the medium. If, now, A is transmitting to Device B that is in transmission range of both A and C, a transmission of C would interfere at B with the transmission from A. This is commonly known as the hidden terminal problem.

On the other hand, if B is transmitting to A and C wants to transmit to D, C senses the medium as used although the transmission would not interfere with the transmission of B at A, since A is not in transmission range of C. This is commonly known as the exposed terminal problem.

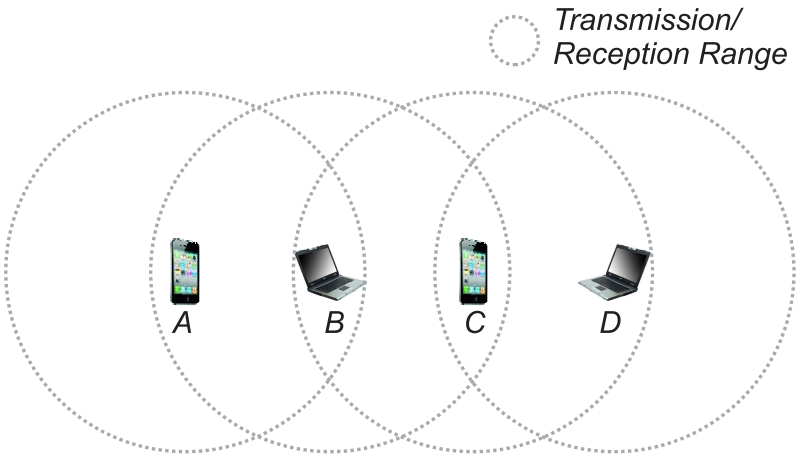


Figure 2.1: Exemplary topology of a wireless multi-hop network where the hidden and exposed terminal problems may occur

A possible way to mitigate these problems is to use a two-way handshake of sender and receiver. Before starting a data transmission, the sender sends a request to send to the receiver. The request to send contains an estimate for the time, the transmission will take. The receiver replies with a clear to send which also contains the estimated transmission time.

In our example for the hidden terminal problem, if B sends the clear to send to A before A starts its transmission to B, C will also receive the clear to send. From this, C knows that the medium is busy at B and that its transmission would interfere with the transmission received at B. This way, an interference at B due to the hidden terminal problem is avoided.

On the other hand, if B transmits to A and C receives the request to send but not the clear to send, C knows that its transmission will not interfere with the transmission of B at A. Thus, C may start its transmission to D in parallel to the transmission of B to A which avoids the exposed terminal problem. For more detailed information, we refer to [104].

An important implication of the request to send/clear to send medium access control mechanism is that it results in symmetric wireless links which is the basis for most intrusion detection schemes in mobile ad hoc networks as we will discuss hereinafter.

More sophisticated medium access control schemes for other types of infrastructure-less networks like sensor or mesh networks exist. However, these are not of relevance for the contribution of this thesis.

2.2.3 Network Layer

A central challenge regarding the network layer of multi-hop mobile ad hoc networks is to discover and maintain routes in the potentially constantly changing network topology. A large family of routing protocols tailored to the infrastructure-less nature were proposed in the last years. For our needs, these can be categorized into proactive, reactive, and geographic approaches.

Proactive protocols exchange topology information constantly, between all nodes of the network. Based on this information, routes to arbitrary nodes can be calculated and, if the network is not partitioned, are readily available when needed. On the downside, exchanging topology information constantly causes a permanent message overhead that increases along with the network size.

Reactive protocols exchange topology information not unless a route is required. In this case, route request messages are broadcasted in the network. Based on the path this broadcast follows to reach the requested node, unicast routes can be established. Since routing information is exchanged only when a route is required, the overhead can be decreased compared to proactive protocols, in particular, in large-scale networks with sparse communication.

Geographic routing protocols aim at further limiting the message overhead generated for routing tasks. This is achieved by restricting broadcast messages to a geographic area where the destination is assumed to be located. On the downside, geographic approaches may fail if (large) obstacles are located between source and destination, thus blocking the area to which broadcast messages are restricted.

The location-based intrusion response mechanism developed in this thesis can be adapted to proactive, reactive, and geographic protocols. Yet, we decided to base the evaluation upon the reactive ad-hoc on-demand distance vector (AODV) protocol [77] described below. We made this choice in light of the large-scale application scenario we outlined in Chapter 1. Here, a proactive protocol would cause a large message overhead also if sophisticated routing protocols that combine proactive and reactive mechanisms to limit the overhead are deployed [119]. We decided not to use a geographic routing protocol since this would have side effects on the metrics we use to evaluate the location-based intrusion response that, itself, introduces aspects of geographic routing.

AODV

We use an implementation of AODV which meets the specifications of RFC3561 [78].

In order to illustrate the details of AODV's mode of operation that are of relevance for our work, we assume that a source node S wants to send data to a destination node D for which no route is known at S. AODV is a reactive routing protocol, which means that a route discovery process is initiated not unless a route is needed. To discover a route, S generates a route request message $rreq_{SD}$. Among other

information, this message contains the network addresses ip_S and ip_D of source and destination, a hop count hop_S, a destination sequence number sn_{dest}, and an originator sequence number sn_{orig}.

A sequence number sn_X is maintained by every node X. For our needs, it is sufficient to say that sequence numbers are used to indicate the age of routing messages. X increments its own sequence number sn_X by one every time it initiates a route request message and uses sn_X as sn_{orig}. Since, in our example, S has no valid route to D, it has no knowledge about the current sequence number sn_D of D. This is indicated by a corresponding flag in $rreq_{SD}$.

The route request message $rreq_{SD}$ is sent as a broadcast. When a node $Y \neq S$ receives $rreq_{SD}$ via a node $X \notin \{Y, D\}$, Y checks whether a change of its routing table is necessary. A corresponding entry in Y's routing table states that it can reach node S via node X. The entry contains (besides other information) the network address of S, the sequence number of S, sn_S, which is included in $rreq_{SD}$ as sn_{orig}, and the distance in hops to S which is included in $rreq_{SD}$ as hop_S. A new entry is created if no entry for S exists in Y's routing table. An update is performed if an entry for S exists with sn'_S and hop'_S, and $sn_S > sn'_S$ or $sn_S = sn'_S$ and $hop_S < hop'_S$, that is, if the route is newer or shorter. After performing the updates to the routing table, Y increments hop_S by one and forwards $rreq_{SD}$.

Since sending route requests as broadcast messages causes a high network load, AODV may use an expanding ring search to mitigate this effect. In this process, routes are first searched in the neighborhood of the source. For this, the time to live (TTL) of RREQ messages is set appropriately. [78] proposes to search routes consecutively in the 1-, 2-, 3-, 5-, and 7-hop neighborhood of the source. If a route is not found in these steps, the search is extended to the full network diameter which is assumed to be 35 hops in [78].

When $rreq_{SD}$ arrives at the destination node D, D generates a route reply message $rrep_{DS}$. Besides other information, $rrep_{DS}$ contains the network addresses of S and D, a hop count hop_D, and the destination sequence number sn_{dest}. In our example, D uses its current sequence number sn_D as sn_{dest}. $rrep_{DS}$ is sent to S as a unicast message, which is forwarded based on the routing table entries that have been created while $rreq_{SD}$ traversed the network. Upon receiving $rrep_{DS}$ at a node $X \neq D$, X updates the entry for D in its routing table in analogy to the update that is performed at X for node S upon receiving $rreq_{SD}$. That is, a route is updated if it is newer or shorter. According to this, X will forward $rrep_{DS}$ only if $rrep_{DS}$ caused X to update its routing table.

The route from S to D may become invalid. This can happen, for example, because of the mobility of nodes due to which a node Y via which a node X forwarded messages to D moves out of X's transmission range. In this case, X notifies the source node S by sending a route error message $rerr_D$ for the destination node D. Upon receiving $rerr_D$, S may initiate a new route request for D. In this case, the expanding ring search does not start from the very beginning, but from the ring in which the destination was found previously.

2.3 Related Work on Attacks on Mobile Ad Hoc Networks

A mobile ad hoc network operates based on the cooperation of nodes that are usually not within the administrative domain of a service provider, but controlled directly by the end user. This makes it easy for an attacker to tamper with the behavior of a node in order to launch an attack on the network. Thus, attack possibilities in mobile ad hoc networks are manifold and exist on each layer of the 5-layer communication model. A comprehensive review can be found, for example, in [115]. Two mechanisms that, in combination, form a very effective denial of service attack are the black hole and the Sybil attack.

The black hole attack [115] can be classified as an active attack on the network/routing layer. From an abstract point of view, a black hole in a mobile ad hoc network is comparable to a black hole in astronomy. The latter attracts matter and electromagnetic waves (physicists may debate this definition), while the first attracts data. In both cases, escape from the black hole is not possible. In the networking context, this means that data which has been attracted by a black hole attacker is dropped instead of being forwarded to the intended receiver. Further continuing with our analogy, false routing information

that is injected into a network corresponds to the gravity of the astronomic black hole. With this, a black hole attacker in a mobile ad hoc network achieves an attracting effect, in that it offers performant routes in terms of the metrics used by the particular routing protocol. As a result, the black hole is chosen to be part of a large number of routes. The negative effects of a black hole attack on the reliability and performance of a mobile ad hoc network are studied, for example, in [37, 1]. The detailed mode of operation of a black hole attack depends on the particular routing protocol deployed. The specific behavior used in our studies will be described in Section 3.1.

The concept of the Sybil attack goes back to [92], a book about a woman who suffered from dissociative identity disorder and developed sixteen different personalities during the course of the disease. By taking this analogy, an attacker that performs a Sybil attack within a networking context presents itself under different network identities. The purposes of a Sybil attack vary from consuming more network resources than intended for one device/user to the prevention of being detected, traced back, or penalized by security mechanisms. A Sybil attack can be performed on every layer of the 5-layer model that deals with identities or addresses. The particular form used in our experiments will be described in Section 3.1.

2.4 Related Work on Intrusion Prevention in Mobile Ad Hoc Networks

Preventive security mechanisms such as gateways or firewalls used in wired and infrastructure-based networks can hardly be applied in mobile ad hoc networks. Thus, preventive security mechanisms in the context of a mobile ad hoc network mainly appear in the form of secure routing protocols. The goal of these protocols is to prevent false routing information from being injected into the network and to avoid the alteration of correct routing information by an adversary. Since the first routing protocols for mobile ad hoc networks were not designed for security but rather for functionality, injecting false routing information or altering correct routing information was possible with little effort. This weakness is exploited for instance by the black hole attack. An overview of secure routing protocols for mobile ad hoc networks can be found in [38]. In general, the functionality of secure routing protocols is based on the deployment of cryptographic measures in order to authenticate routing information.

As a precondition, mechanisms for key establishment and for authenticating routing information used during route discovery are required. A mechanism for key establishment in mobile ad hoc networks without a central, trusted entity can be found, for example, in [83]. Authenticating routing information is, in general, based on hash chains. Here, a commonly known secure hash function is applied multiple times to a seeding value. The result can be included in a data packet along with the seed, which is hashed at each hop. Using, for example, the maximum time to live of a packet to generate the result of the hash chain, the correctness of the hop count can be verified at each node. To authenticate the source of a broadcast, the values of a hash chain can be used as one-time keys to produce message authentication codes. If keys are released delayed with respect to the message they were used to authenticate, the authenticity of the sender can be verified. One of the first protocols authenticating broadcast sources based on hash chains is presented in [79]. Subsequent approaches, such as [43, 112] are designed to overcome limitations such as relatively high authentication delays and overhead caused in scenarios with many broadcast sources.

Two specific secure routing protocols are SAODV [117] and Ariadne [39]. These protocols were designed as secure versions of AODV [77] and DSR [48], respectively. Despite being designed carefully, these protocols were proven to be susceptible to attacks, recently. New attack mechanisms are identified in [40, 2, 3]. To cope with these, new secure routing protocols and frameworks to mathematically prove their security are proposed in [2, 3]. Developers of secure routing protocols became cautious about claiming perfect security of their protocols. Instead, well defined attacker models specify the types of malicious behavior which can be prevented. Furthermore, secure routing protocols are not able to thwart attacks on other layers such as jamming attacks on physical layer or attacks on wireless medium access control protocols. Also, some of the proposed approaches introduce a considerable computa-

tional and communication overhead that causes negative effects on network performance even when no misbehaving devices are present.

2.5 Related Work on Intrusion Detection and Intrusion Response in Mobile Ad Hoc Networks

Intrusion detection systems in combination with intrusion response mechanisms can be deployed as reactive security measures to establish a second line of defense for the case of subverted preventive security measures. Intrusion detection systems are able to identify ongoing misbehavior. For this, two different approaches, anomaly-based and signature-based detection, can be deployed. Anomaly-based intrusion detection systems 'know' the normal operational bounds of a network and can classify any deviation from this as an attack. Signature-based intrusion detection systems 'know' the characteristic patterns (signatures) of attack mechanisms and can recognize these. As an example, we will consider the detection of a black hole attack. In this case, anomaly-based intrusion detection can be based, for example, on the measurement of the packet loss in a network without misbehaving nodes. If the loss increases, an anomaly-based system can classify this as a potential attack.

For signature-based detection of a black hole attack, a pattern for this attack behavior has to be specified. The pattern depends on the particular routing protocol deployed and describes individual facets of the attack like the injection of false routing information. Both anomaly-based and signature-based approaches have advantages and disadvantages. However, since our goal is not to design a novel intrusion detection system for mobile ad hoc networks, a detailed evaluation is beyond the scope of this thesis.

Similar to preventive security measures, intrusion detection systems can not be transferred directly from wired and infrastructure-based environments. The negative effects of ad hoc routing protocols on intrusion detection systems that were designed for wired and infrastructure-based networks are shown in [4]. Thus, new approaches for intrusion detection in mobile ad hoc networks were developed in the last years.

One of the first approaches for intrusion detection and response in mobile ad hoc networks is presented in [65]. This anomaly-based system uses a packet loss metric to detect misbehaving nodes. For this, an intrusion detection component (the watchdog) on each node monitors the forwarding behavior of its neighbors. If a certain threshold of packet loss is exceeded, the watchdog sends a notification to the source of the packets. Intrusion response is performed by the pathrater component, which collects information from the watchdogs, calculates ratings for each node, and sums up these node ratings into a rating for the route to the destination. In case multiple routes to a destination exist, the pathrater chooses the one with the best rating for transmission. By doing so, the pathrater prevents the inclusion of detected malicious nodes in a route.

Subsequent approaches for intrusion detection were developed with the goal of overcoming limitations of the watchdog/pathrater approach and of further taking into account constraints of devices and harnessing the characteristics of mobile ad hoc networks.

The anomaly-based intrusion detection system proposed in [41] uses a clustering approach to reduce the system load caused by intrusion detection tasks. For this, one node out of a set of neighboring nodes (a cluster) is determined as the clusterhead. In each cluster, only the clusterhead is responsible for intrusion detection. The role of being clusterhead is assigned periodically to different nodes in a cluster.

In [118], a cooperative, anomaly-based intrusion detection system is presented. Here, an exchange of intrusion detection information among nodes is used to enhance detection performance. With this, it is possible to lower false positive (a detection when there is no attack) and false negative (no detection when there is an attack) rates.

A signature-based intrusion detection system for the AODV routing protocol is presented in [107]. The signature of an attack is modeled as a composition of the individual steps which it consists of. Using finite state machines, an attack is recognized if the state machine consecutively passed through all states, each of which represents an individual step of an attack.

An anomaly-based approach for detecting colluding misbehaving nodes is presented in [27]. In contrast to the intrusion detection approaches discussed before, the approach is able to deal with hop-by-hop encrypted wireless links.

Comprehensive approaches that take into account intrusion detection as well as intrusion response are, for example, CONFIDANT [13], CORE [67], OCEAN [7], Routeguard [32] and AntSec [71, 72, 70]. For each of these, intrusion detection is performed by an anomaly-based component whose core idea is comparable to the watchdog as proposed in [65]. The information that is gathered by the watchdog component is collected and managed in a reputation system that is either based on locally available information only or global information that is exchanged between nodes. The assignment of reputation values to nodes is based on the addresses of nodes as identifiers in all approaches. Consequently, intrusion response is the exclusion of misbehaving nodes with a bad reputation from the network based on their addresses. Generally speaking, no data is sent to/via or received from/via addresses that are assigned with a bad reputation.

An approach that is to some extent related to the location-based intrusion response system we develop and evaluate in the following chapters is TIARA [86]. The authors propose a set of intrusion detection and intrusion response mechanisms that can be implemented independently from a particular routing protocol. One part of TIARA describes a distributed wireless firewall in which dynamic filter rules are enforced by each benign node in the network. Malicious traffic is filtered by the nodes surrounding the source. This results in a quarantine effect that is comparable to the quarantined areas that are the fundament for the location-based intrusion response mechanism developed in this thesis. However, firewall rules are based on network addresses and, thus, are susceptible to Sybil attacks the same way as other address-based intrusion response systems. An evaluation of the mechanisms proposed is not presented.

A mechanism for determining geographic abstractions of jammed areas in sensor networks, which correspond to the quarantined areas we introduce in mobile ad hoc networks, is presented in [114]. The focus is set on designing and evaluating a distributed protocol for defining the boundaries of jammed areas properly, when a global information on the location of nodes is not available. Although the approach was designed for sensor networks, the assumptions made do not prevent applying it in mobile ad hoc networks. Thus, the approach presented can be used as basis for establishing quarantined areas in mobile ad hoc networks without global knowledge on the location of nodes.

In [35], a model for the spreading of mobile worms via short distance wireless links of cell phones is developed. The model is used to calculate quarantine boundaries at which immediate countermeasures can be taken to stop the spreading. Due to the assumptions made for the network model, the work is not applicable in our context.

2.6 Related Work on Geographic Routing Protocols for Mobile Ad Hoc Networks

Several routing mechanisms for mobile ad hoc networks that take into account geographical information of nodes were proposed. An overview can be found, for example, in [66] or [59]. For our needs, it is sufficient to understand the basic idea of geographic routing. Thus, we shortly introduce LAR [54] and DREAM [9], two of the first protocols which are related to our work. Both protocols use location information to restrict the propagation of broadcast messages in order to reduce routing overhead.

LAR proposes two schemes to improve the overhead of the route discovery phase of on-demand routing protocols. For the first scheme, a request zone is specified such that it contains the initiator of a route request and an area in which the destination is expected to be located. Route request messages are only forwarded by nodes that are situated within the request zone. The second scheme is based on the distance between the destination and nodes that could potentially forward a route request. Here, a node only forwards a route request if it is closer (within certain bounds) to the destination than the node it received the request from.

DREAM was developed as a robust protocol that sends every message (not only route requests) as a restricted broadcast. To restrict the broadcast, DREAM determines the direction from sender to receiver based on their geographical positions. Only nodes that are (within certain bounds) located in this direction forward messages. At this point, it is necessary to mention that the objective of DREAM was the efficient dissemination of the location information of nodes throughout a network. However, a detailed description of this process is beyond the scope of this thesis.

A precondition for LAR and DREAM as well as for our approach is that nodes are aware of their geographical position. One way to determine this would be the use of GPS [22]. Besides this, other approaches for determining positions in dynamic environments were proposed. A survey can be found in [99]. The localization mechanisms presented there were developed for sensor networks but can also be applied to mobile ad hoc networks. Localization techniques include beacon based approaches, in which landmarks emit position information that can be processed (for example, by triangulation) by mobile devices. For indoor positioning, badges can be used to support position determination on the granularity of rooms if triangulation is not feasible. Besides these mechanisms, which in general allow for determining a globally unique position, approaches were proposed that enable relative positioning of nodes to each other. Since our approach for location-based intrusion response neither requires globally available information nor globally unique positioning, most of the localization approaches proposed can be deployed.

2.7 Related Work on Analytical Models for Mobile Ad Hoc Networks

A metric often described analytically is the throughput of mobile ad hoc networks. One of the first models for this is presented in [30]. The authors develop an upper bound for per node throughput subject to the total number of nodes in the network and subject to the theoretical transmission speed of devices. [30] served as basis for several further studies on the throughput of mobile ad hoc networks. In [49], the influence of the transmission range on the upper bound proposed in [30] is described. An analysis of the effect of different traffic patterns is performed in [60]. The impact of node mobility is described in [29]. In [20], network coding is considered as a means to increase per node throughput.

The end-to-end delay in mobile ad hoc networks subject to the number of nodes and to node mobility is modeled in [28]. Closely related to the delay are the length (geographical distance of sender and receiver) and the lifetime (time until nodes move out of each others transmission range) of single-hop links and multi-hop routes. A model for the length of single-hop links for a uniform as well as for a Gaussian distribution of node locations is proposed in [69]. The route length for a uniform node distribution is described in [36]. The lifetime of single-hop links subject to node mobility is analyzed in [89]. In [5], analytical models for the lifetime of multi-hop routes are developed for different mobility models. In [108] the lifetime of single-hop routes is modeled for different mobility patterns.

Recently, game theory found its way into research on mobile ad hoc networks. From an abstract point of view, game theory is a tool for modeling give-and-take interactions of two parties. In [68], an analysis of the CORE intrusion detection and response system [67] based on game theory is presented. In [100], the authors propose game theoretic views on issues such as transmission power control, medium access strategies, and packet forwarding. A model based on game theory for misbehavior in mobile ad hoc networks is the focus of [106]. Closely related to this is the work in [46]. Here, a reputation mechanism used to decide on whether to interact with a node or not (that is, relaying its packets or using it as a relay) is developed based on game theory. Since our work is not aimed at describing the interaction of two particular nodes, we did not choose game theory as basis for our models.

To the best of our knowledge, [36] is the only work that describes the effects of node misbehavior in mobile ad hoc networks by means of a geographical model. However, it focuses on the route-length distribution as a metric only and does neither include modeling of packet loss nor of security mechanisms.

2.8 Related Work on Adaptive Transmission Power in Mobile Ad Hoc Networks

Adaptive transmission power has a wide range of applications in wireless networks. Saving battery power or signal strength control for CDMA-based systems are prominent examples. In mobile ad hoc networks, adaptive transmission power is mainly used for controlling and optimizing the network topology. Algorithms differ in their constraints (for example, node degree, throughput) subject to which an optimal transmission power shall be achieved. A further distinction can be made regarding a network-global or a per node adaptation of the transmission power (that is, whether all nodes use the same transmission power or each node uses an individual value). Also the overhead generated (in terms of messages) to coordinate the use of an adaptive transmission power between individual nodes categorizes the existing algorithms.

One of the first approaches of power aware routing in mobile ad hoc networks is proposed in [97]. Metrics for optimal routing with respect to energy consumed are specified and validated by simulation. Yet, a direct adaptation of the transmission power is not considered. In previous and subsequent works, the authors also address energy awareness for medium access and transport layer protocols not specifically related to mobile ad hoc networks.

Distributed heuristics for topology control without the necessity to exchange additional control information are proposed in [85]. The algorithms are based on information that is available by routing protocols. Thus, no additional information has to be exchanged. Power is adapted on a per-node basis. Variations for link state and distance vector protocols are described.

A distributed protocol for topology control in order to achieve a connected network by adapting transmission power such that an optimal number of neighbors per node is maintained is proposed in [10]. In contrast to [85], control messages are needed.

Algorithms for adaptive network-global as well as individual transmission power in mobile ad hoc networks with the goal of achieving a maximized throughput (not minimal energy) subject to the network load and the network density (nodes per area) are proposed in [75]. No additional messages are needed, but, for individual transmission power, the IEEE 802.11 protocol is extended to prevent asymmetry. An analytical model for the algorithms is proposed in [76]. A protocol that can be used as an addition to other (non-power-aware routing protocols) is proposed in [25]. The protocol redirects communication between adjacent nodes by introducing additional forwarding nodes. This way, the overall transmission power required in a route is minimized. Control messages are needed to manage the redirection. Analytical models for the comparison of static and adaptive transmission power considering mobility during route discovery as well as route maintenance phases have been formulated and analyzed in [24].

A protocol for maximizing network lifetime by adaptive transmission power on a per-node basis is proposed in [98]. The algorithm works in a distributed way and, for this, requires the exchange of corresponding protocol messages.

To the best of our knowledge, no related work exists that utilizes an adaptive transmission power in the context of location-based security mechanisms for mobile ad hoc networks.

2.9 Related Work on Delay Tolerant Communication in Mobile Ad Hoc Networks

Several approaches to enable delay (or disruption) tolerant communication were proposed in literature. Out of these, we discuss performance enhancing proxies [12], the Licklider Transmission Protocol [84], and the Bundle Protocol [94] for potential deployment in our scenario.

Performance enhancing proxies, in general, split up connections into different sections each of which is optimized for the individual conditions of the corresponding network segment. Implementations exist on transport and application layer of the 5-layer model. A well known example for a performance enhancing proxy on application layer is a proxy server for web browsing. Examples for proxied variants of TCP include snoop TCP [6] and Freeze-TCP [23]. TCP proxies adapt TCP to the conditions in wireless environments in order to deal with lossy wireless channels and periodic disconnections.

The Licklider Transmission Protocol and the Bundle Protocol were proposed by IRTF's Delay Tolerant Networking Research Group to enable efficient communication in disruptive environments as, for example, interplanetary networks. The Licklider Protocol is implemented above the data link layer and, thus, is designed to operate over one-hop wireless links with very high delays in the scale of minutes up to hours. The Bundle Protocol works on an additional layer between transport and application layer. In a store-and-forward manner, it takes data from application layer and uses common transport protocols like TCP or UDP to transmit the data (possibly over multiple hops) to the receiver intended, as soon as the next hop in this direction is reachable. Together, the Licklider Protocol and the Bundle Protocol form a delay tolerant network stack, where the Licklider Protocol is responsible for a reliable communication between adjacent nodes, and the Bundle Protocol provides end-to-end transmission.

Security issues of delay tolerant networks are studied in [14]. The authors scrutinize the inherent resistance to attacks that comes along with disruption tolerant networking techniques.

3 Location-based Intrusion Response in Mobile Ad Hoc Networks

In this chapter, we present the basic version of location-based intrusion response in mobile ad hoc networks. Figure 3.1 shows a schematic representation of the mode of operation. By establishing quarantined areas void of communication at locations where misbehavior was detected, we exclude misbehaving nodes from the network. This way, the functionality of the intrusion response mechanism does not depend on network addresses, which can be changed with little effort in an environment that is beyond a central control.

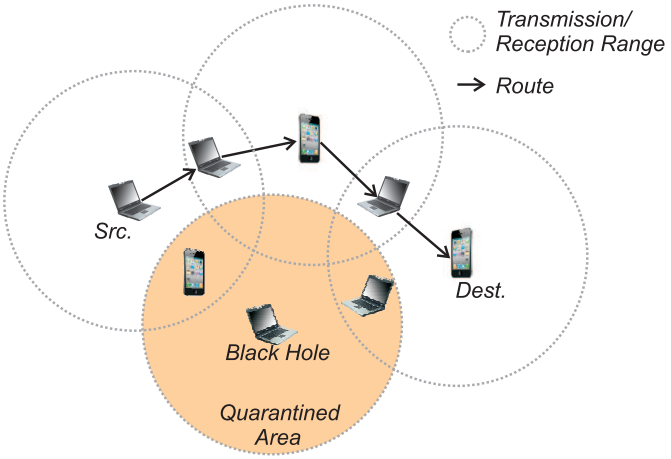


Figure 3.1: Schematic representation of the location-based intrusion response. Please note that for reasons of clarity the transmission/reception range is not shown for each device.

We compare the performance of the location-based intrusion response to an address-based approach using the example of a combined black hole and Sybil attack on a mobile ad hoc network. The comparison is based on a series of simulation studies. For key metrics describing the performance of the location-based approach, such as the packet loss that is caused by black holes, we further present an analytical model and mutually validate model predictions and simulation results. We start this chapter with a description of the relevant details of the system architecture.

3.1 Architecture

In this section, we present details of the design of the attacks, the location service, the intrusion detection system, and the address-based and the location-based intrusion response mechanisms.

3.1.1 The Black Hole Attack

We realized the black hole behavior outlined in Section 2.3 for the routing protocol AODV described in Section 2.2.3. The detailed mode of operation of a black hole B can be summarized as follows:

- B answers each route request message $rreq_{SD}$ with $B \neq S$ it receives.
- B claims that the intended destination D is its direct neighbor, that is, B generates a route reply message $rrep_{DS}$ with $hop_D = 1$. This is the first step towards achieving the attracting effect as described in Section 2.3. If the route that is offered by B appears to be shorter (measured in hops) than the route offered by D, B's route will be chosen by S according to the specifications of AODV.

- B increments its local sequence number sn_D for the intended destination D by one every time it receives a route request $rreq_{SD}$. B uses this sequence number as sn_{dest} when it generates $rrep_{DS}$. This is the second step towards achieving the attracting effect. From route requests $rreq_{DX}$ that are issued by D for an arbitrary destination X, B learns the current sequence number of D. With this sequence number, D will answer incoming route requests. Since B increments this sequence number when answering a route request on behalf of D, the route offered by B will appear to be newer than the route offered by D. Thus, the route offered by B will be chosen by S.
- B does not forward packets.

3.1.2 The Sybil Attack

To achieve a Sybil behavior, we extended the definition of a black hole B described above as follows:

- B changes its network layer address $ip_{B,n}$ at a Sybil frequency f_{syb}.
- B keeps track of its previous addresses

  $$\left\{ip_{B,1}, ..., ip_{B,n-1}\right\}$$

  With this, B is still able to handle (that is, drop) messages that it receives due to a previously generated route reply for which B used the address

  $$ip_{B,x} \in \left\{ip_{B,1}, ..., ip_{B,n-1}\right\}$$

- B chooses its addresses $ip_{B,1}, ..., ip_{B,n}$ such that no address appears more than once during a simulation to avoid address conflicts. That is, at the end of the simulation

  $$\left\{ip_{B,1}, ..., ip_{B,n}\right\} \cap \left\{ip_{X,1}, ..., ip_{X,n}\right\} = \emptyset$$

  applies for an arbitrary node $X \neq B$.

3.1.3 The Location Service

To determine the location of nodes, we implemented a location service that offers each node the possibility to ask for the location of any node in the network. The service is based on the global and exact knowledge of positions we have in our simulation tool.

The globally available and globally unique position information as it is provided by the location service is not a precondition for the location-based intrusion response, but was chosen to reduce the complexity of the implementation. The location-based approach can operate on location information that is only available locally, in the neighborhood of a misbehaving node, since quarantined areas are managed on a per node basis and not globally for the network. We describe this in more detail in Section 3.1.6. Thus, all of the localization approaches we outlined in Section 2.6 could be deployed in a real-world implementation.

3.1.4 The Intrusion Detection System

For detection of the combined black hole/Sybil attack, we use an anomaly-based approach. Detection is based on the rate at which a node drops packets that should have been forwarded. Our system is comparable to the one presented in [118], albeit it works on quite an abstract level. Here it is important to note that our goal is not the design and evaluation of a new intrusion detection system, but of the location-based intrusion response mechanism. Thus, we implemented an intrusion detection system that makes use of the bird's eye view that is available in our simulation tool. However, the intrusion detection system is tunable to adjust detection performance (detection speed, true/false positives/negatives) to that of real-world intrusion detection systems. For this, our intrusion detection system works in the following steps.

- Monitoring: During a monitoring interval t_{mon}, a node X collects information about the forwarding characteristics of its neighbors. We call a node Y a neighbor of a node X if X is within the transmission range of Y. For each of its neighbors, X maintains a counter $n_{Y,rec}$ for packets that Y received from another neighbor of X or from X itself for forwarding. A second counter $n_{Y,forw}$ is maintained for packets that Y forwarded to another neighbor Z or to X itself. Regarding the abstraction mentioned above, monitoring is not implemented in detail on the medium access layer of nodes as it is necessary in a real system, but based on the globally available knowledge about the forwarding behavior of nodes in the simulation environment.
- Classification: After each monitoring interval t_{mon}, X calculates a rating $r_{X,Y}$ that describes the forwarding behavior for each of its neighbors Y. We use a weighting factor $w_{bal} \geq 1$ to balance the counters $n_{Y,rec}$ and $n_{Y,forw}$. This way, we prevent false positives if packets are dropped due to inherent network conditions such as collisions or exhausted queuing space. w_{bal} packets that are forwarded correctly outbalance one dropped packet. With this, $r_{X,Y}$ is calculated as

$$r_{X,Y} = \max\left(n_{Y,rec} - \frac{n_{Y,forw}}{w_{bal}} + r'_{X,Y}, r_{max}\right)$$

  where $r'_{X,Y}$ denotes the previous rating that X calculated for Y and r_{max} denotes a maximum node rating used to prevent that benign nodes are excluded permanently from the network due to repeated false positive detections. If $r_{X,Y}$ exceeds a certain threshold value $thres_{black}$, X classifies Y as a black hole.

Similar to OCEAN [7], our intrusion detection system works on a per node basis. No information about the calculated ratings is exchanged between nodes. Although detection performance can be increased when nodes act cooperatively, the evaluation of OCEAN showed that feasible results can be achieved when only local information is considered. Also for a real system we would preferably use a local detection, since it avoids the overhead of exchanging and authenticating intrusion detection information.

3.1.5 The Address-based Intrusion Response Mechanism

We designed the address-based intrusion response strategy such that a node Y, which was classified as a black hole by a node X based on the rating $r_{X,Y}$, is excluded from the network based on its (current) network layer address $ip_{Y,n}$. For this, X does not send/accept any packets to/from $ip_{Y,n}$. If X is part of a route from a Source S to a Destination D, for which Y was X's next hop, X will inform S by sending a corresponding route error message $rerr_D$.

3.1.6 The Location-based Intrusion Response Mechanism

In contrast to the address-based intrusion response mechanisms, the location-based approach excludes an identified black hole based on its location. If a node B is classified as a black hole by a node X based on the rating $r_{X,B}$, X obtains B's location from the location service and establishes a quarantined area with radius d_{quar} at B's location. If X is situated within the quarantined area, it will not forward any messages. If, in this case, X is part of a route from a source S to a destination D, X will inform S by sending a corresponding route error message $rerr_D$.

Since X will not forward any messages while it is quarantined, route request messages that are usually sent as broadcast messages will not reach the black hole B, as shown in Figure 3.2. Thus, we prevent B from being a part of newly established routes while it is quarantined. This effect is comparable to the restriction of broadcast messages as used by the location-based routing protocols LAR [54] and DREAM [9].

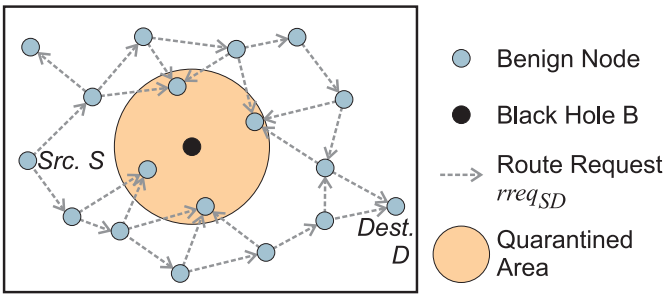


Figure 3.2: Restricted broadcast of Route Request Messages in the location-based intrusion response

We assume that a tracking of B is not possible within a quarantined area. Therefore, an adaptation of the quarantined area when B moves is not possible. For this reason, a quarantined area is revoked after the period t_{reset}. If B leaves the quarantined area, it may again become part of routes and drop packets. B is then classified as a black hole again and a new quarantined area is established.

Although we chose AODV as basis for our evaluation, the location-based intrusion response is neither bound to AODV in particular, nor to reactive routing protocols in common. Proactive routing protocols are, in general, based on constantly disseminating topology information in the network. From this, nodes can deduce global topology information based on which routes can be calculated. When used together with proactive protocols, instead of preventing routing information from entering a quarantined area as it is done for AODV, the location-based approach would prevent proactive routing information from leaving a quarantined area. As a result, falsified routing information as it would be generated by a black hole node is restricted to a quarantined area and kept away from other parts of the network.

3.1.7 Implementation Details

We integrated the components described above in the JiST/MobNet simulation tool as presented in [56]. JiST/MobNet is based on the JiST/SWANS simulation tool [8] tailored to mobile ad hoc networks, which was reorganized and extended at our institute. Besides other functionality, we added the attacks and security mechanisms described previously.

Figure 3.3 depicts the interaction of the different components schematically. We implemented the attacks and the intrusion response mechanisms in a realistic way, as we would for deployment in a real-world mobile ad hoc network. To keep the implementation effort on an affordable level, the location service and the intrusion detection system are implemented based on the bird's eye view that is available in the simulation tool.

3.2 Evaluation

In this section, we present the evaluation of the location-based approach for intrusion response. For this, we follow the methodology proposed by Jain [45].

3.2.1 Goals of the Evaluation

Goal of the evaluation is to compare the performance of the location-based intrusion response with an address-based approach when confronted with a combined black hole/Sybil attack as described in the previous section. We state to what extend the intrusion response mechanisms are able to recover network functionality. As baseline, we use the performance of a network without misbehaving nodes.

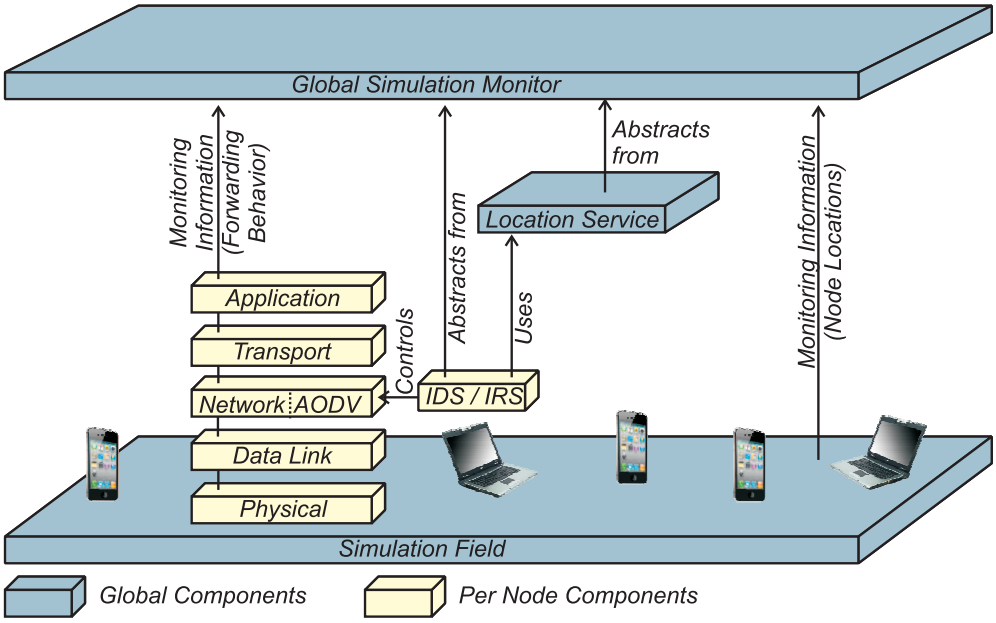


Figure 3.3: Interaction of the components in the simulation tool

3.2.2 Services of the System

The system evaluated is a mobile ad hoc network built upon the following protocols.

- Physical and data link layer: IEEE 802.11 DCF based on the request to send/clear to send handshake to coordinate medium access as described in Section 2.2.2. Thus, bidirectional wireless links can be assumed for intrusion detection.
- Network layer: IPv4 with AODV routing service and intrusion detection as well as address-based and location-based intrusion response as described in Section 3.1.
- Transport layer: UDP

In addition, to determine the location of nodes, the location service described in Section 3.1.3 is available.

3.2.3 Metrics for the Evaluation

As metrics to evaluate the performance of the intrusion response mechanisms, we use

- the packet loss caused by black holes as metric which eliminates all side effects that cannot be directly associated with the black hole attack,
- the inherent packet loss of the network, which, as further metrics, includes (1) the packet loss caused by collisions on the wireless medium which may occur despite the request to send/clear to send handshake in case two nodes start transmitting simultaneously, (2) the packet loss caused by exhausted queuing space in the network interface, and (3) the packet loss caused by the AODV protocol due to route breaks or unavailable destinations,
- the packet loss caused by intrusion response, which is the sum of the packet loss caused by a route break triggered by intrusion response and, for the case of location-based intrusion response, the packet loss caused if source or destination are quarantined when data has to be transmitted,
- the performance of the intrusion response system in terms of (1) the detection rate measured in correct detections per monitoring interval in relation to the number of all misbehaving nodes in radio range of the monitoring node and (2) the false positive rate measured in detections when there is no attack in relation to correct detections per monitoring interval,

- the overall packet loss in the network as sum of all loss metrics described above,
- the overall packet delivery ratio to double-check for unrecognized packet loss,
- the average route length measured in hops to identify partitioning effects in the network,
- the transmission delay from sending application to receiving application consisting of (1) the routing delay on the source node and (2) the propagation delay of the network, to quantify the temporal side-effects of quarantined areas that have to be bypassed.

3.2.4 Parameters of the System

A multitude of parameters affect the behavior and the performance of the mobile ad hoc network that is the system under test. For the network itself, these parameters include basic settings such as transmission power of single nodes, antenna characteristics, traffic patterns, or mobility models, which we will specify in the experimental design described below. Additional parameters for the Sybil attack, the location service, the intrusion detection system, and the intrusion response mechanisms are set as introduced in Section 3.1.

3.2.5 Selection of the Factors for the Evaluation

From the large number of parameters that characterize the network, we choose the following set of factors that affect network performance subject to the metrics described above significantly. We consider

- the number of malicious nodes in the network,
- the Sybil frequency f_{syb} at which a malicious node changes its network layer address,
- the interval t_{reset} after which quarantined areas are revoked.

3.2.6 Evaluation Technique

We evaluate the performance of the location-based intrusion response by means of simulation. For this, we use the simulation tool described in Section 3.1.7.

To improve runtime, we use a Condor cluster [105] to distribute simulation runs on multiple processors.

3.2.7 Workload of the System

We chose the parameters and factors described above such that we achieve a moderately loaded network, that is, the network operates in a non-congested state. In light of our application scenario outlined in Chapter 1, we selected mobility of the nodes to correspond to pedestrian speed. With this, we achieve a reasonable lifetime of routes and, thus, a feasible overhead of control traffic subject to data traffic.

3.2.8 Experimental Design

Based on preliminary simulations, we defined the parameters and factors used for the evaluation such that the system operates within normal bounds.

Parameters

The basic parameters of the network are shown in Table 3.1.

Table 3.1: Basic network parameters as used in the simulation studies

Number of nodes	1000
Antenna gain G	$0dB$
Transmission power P_S	$7dB$
Reception threshold $P_{R,min}$	$-81dB$
Signal wavelength λ	$2.4GHz$
Resulting transmission range	$\approx 250m$
Network density	≈ 7.5 neighbors per node in order to typically achieve a connected network
Resulting simulation area	Square with $4750m$ side length
Traffic pattern	Constant bitrate traffic with streams of 2048 bytes per second split in 4 packets per second. 20 streams run in parallel with a duration of 30 seconds each. Streams are set up uniformly within the first 30 seconds of simulated time. After this, every time a stream ends, a new one is created immediately. Source and destination are randomly selected from the set of benign nodes, resulting in that malicious nodes are not chosen as source or destination.
Node Placement	Geometrically uniformly distributed random placement
Mobility model	Random waypoint mobility as described in [48] with minimum speed of 1 meter per second and maximum speed of 2 meters per second. Nodes move continuously; upon arrival at a waypoint, a new waypoint is chosen and movement is continued immediately.

For the intrusion detection system, we chose the parameters shown in Table 3.2, resulting in a detection performance that has the same order of magnitude as the results presented in [118] if all other parameters are defined as described above. For other conditions, the intrusion detection system would have to be adapted appropriately. However, since we focus on evaluating the intrusion response mechanisms, a detailed evaluation of the intrusion detection system is beyond the scope of this thesis. Thus, parameters are chosen such that we obtain reasonable operational bounds of the intrusion detection system, which are the same for the address-based and the location-based intrusion response mechanisms to guarantee equal preconditions for the evaluation.

Table 3.2: Parameters of the intrusion detection system as used in the simulation studies

Monitoring interval t_{mon}	$1s$
Weighting factor $w_{balance}$	1
Rating limit r_{max}	30
Detection threshold $thres_{black}$	10

The radius of quarantined areas for location-based intrusion response is set to $d_{quar} = 250m$, which corresponds to the transmission/reception range of nodes. We decided to keep this parameter fixed for now. A smaller quarantine radius would result in route request messages reaching malicious nodes and is, therefore, not feasible. When we increase the radius of quarantined areas, we get a buffer for mobility of a malicious node as well as for positioning inaccuracy in realistic systems. But, since the quarantined area increases proportionally to the square of the radius, an increased radius could (especially in scenarios with multiple malicious nodes) quickly lead to a partitioned network, which is clearly not our intention. Yet, we will consider a variable radius of quarantined areas together with an adaptive transmission power of devices in Chapter 4.

Factors

The factors considered in the simulation studies are shown in Table 3.3. We performed a full-factorial simulation study in which we simulated each combination for one hour simulated time split up in 6 parts with 10 minutes each to distribute workload and reduce unwanted side-effects of the random waypoint mobility model.

Table 3.3: Factors as used in the simulation studies

Number of misbehaving nodes out of all nodes	1, 2, 3, 5, 10
Sybil frequency f_{syb} in address changes per second	$\frac{1}{60}$, $\frac{1}{45}$, $\frac{1}{30}$, $\frac{1}{20}$, $\frac{1}{15}$, $\frac{1}{10}$, $\frac{1}{5}$, $\frac{1}{3}$, $\frac{1}{2}$, 1
Time t_{reset} in seconds after which quarantined areas are revoked	15, 30, 45, 60, 90, 120, 180, 300, 420, 600

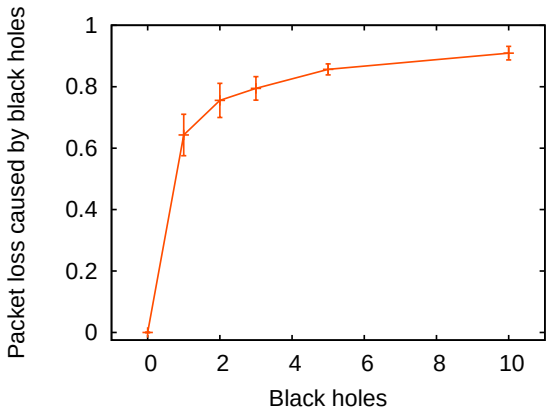
3.2.9 Analysis of the Results

Within this section, we present and interpret the results of our simulation studies. For each metric introduced above, we compare the results (1) for a network without intrusion detection and response subject to the number of misbehaving nodes, (2) for address-based intrusion response subject to the number of misbehaving nodes and subject to the Sybil frequency, and (3) for location-based intrusion response subject to the number of misbehaving nodes and subject to the time after which quarantined areas are revoked. Since the Sybil attack did not show any significant effects on the location-based intrusion response, we do not present the results obtained for the location-based approach subject to the Sybil frequency. All of the following plots are shown with 95% confidence intervals. Note that for reasons of clarity, not all combinations of the full factorial evaluation are presented.

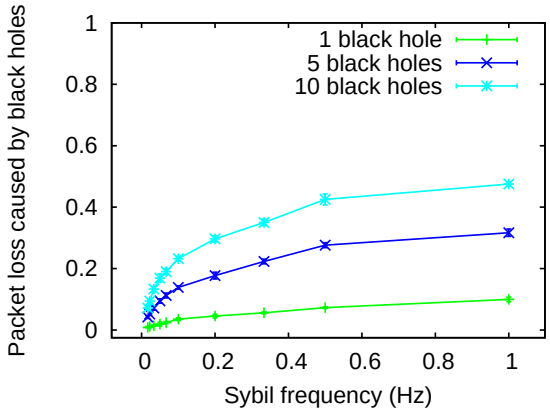
Packet Loss Caused by Black Holes

Figure 3.4a shows the packet loss that is caused by the black hole nodes in a defenseless network subject to the number of black holes. A packet loss of more than 60% for one black hole going up to around 90% for 10 black holes (only 1% of all nodes) illustrates the severe effect of the black hole attack on the availability of the network.

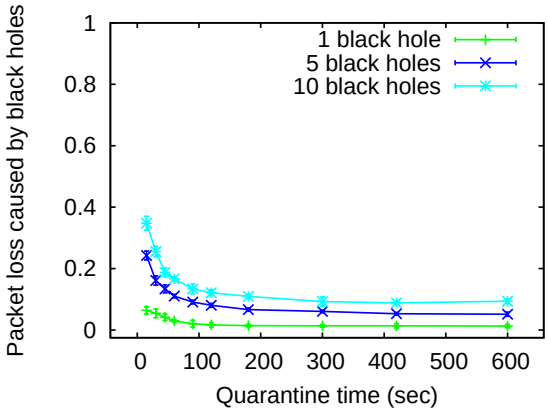
The loss caused by the combined black hole/Sybil attack when the network is defended by address-based or location-based intrusion response is shown in Figures 3.4b and 3.4c, respectively. The address-based solution works efficiently for low Sybil frequencies, that is, if misbehaving nodes do not change addresses. In this case, the loss caused by black hole can be reduced, for example, for the scenario with



(a) Defenseless network



(b) Address-based intrusion response



(c) Location-based intrusion response

Figure 3.4: Packet loss caused by black holes

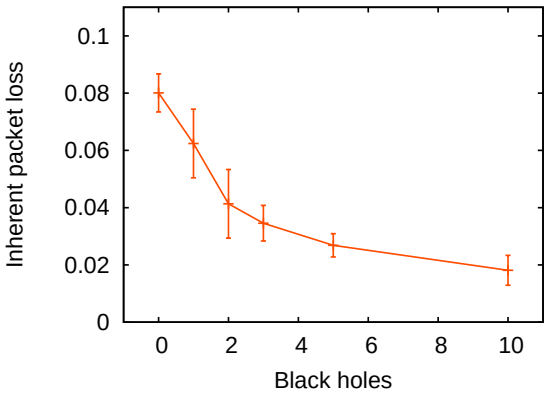
10 black holes to about 10%. As soon as misbehaving nodes increase the Sybil frequency, the protection of the address-based solution decreases. For one address change per second, the loss increases to about 50%.

The performance of the location-based intrusion response depends on appropriate choices of the interval after which quarantined areas are revoked. If this interval is chosen too short, the quarantined area does not protect the new route discovery process after a route affected by misbehavior was interrupted. If the quarantine time is chosen appropriately, the location-based approach achieves a performance comparable to the address-based solution. That is, the loss caused by black holes can be reduced to about 10% in the scenario with 10 black holes.

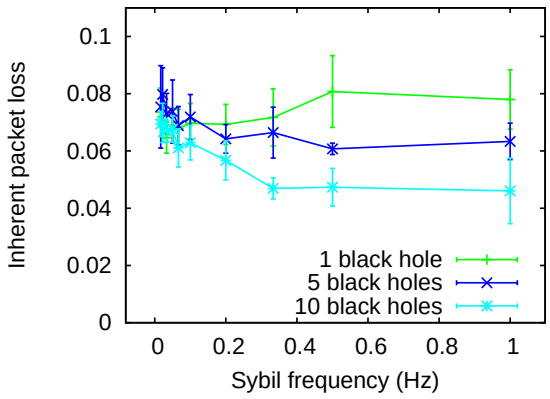
At this point, we have to emphasize that the Sybil frequency is part of the misbehavior and will be optimized by a potential attacker to achieve high loss rates. The quarantine time, on the other hand, is part of the intrusion response and can be controlled in order to recover network performance.

Packet Loss Caused by Inherent Network Properties

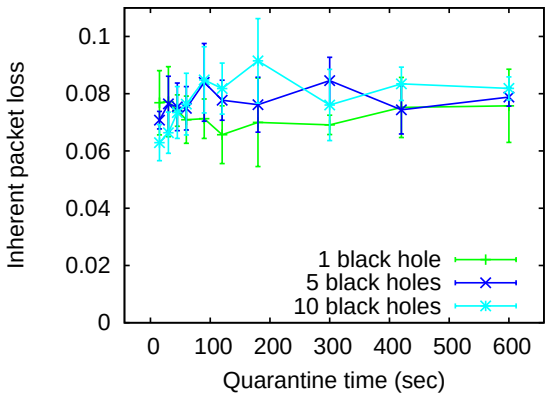
Figure 3.5a shows the inherent loss in a defenseless network caused by collisions on the wireless medium, exhausted queuing space in the network interfaces, and route breaks due to node mobility as well as unreachable destinations. It stands out that the inherent loss decreases significantly if the number of misbehaving nodes increases. To explain this effect, we have to consider the route length subject to the number of misbehaving nodes presented in Figure 3.14a. We see that the length of the routes over which packets can be transmitted successfully decreases if the number of misbehaving nodes increases.



(a) Defenseless network



(b) Address-based intrusion response



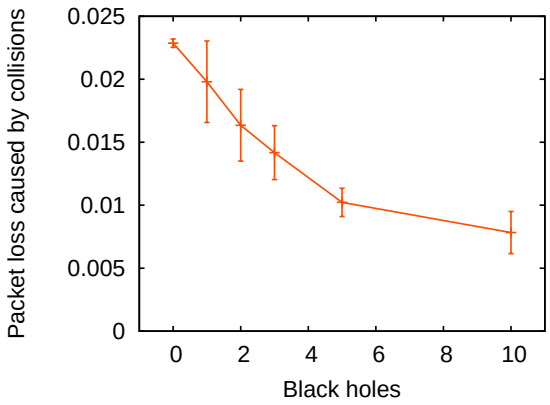
(c) Location-based intrusion response

Figure 3.5: Packet loss caused by inherent network properties

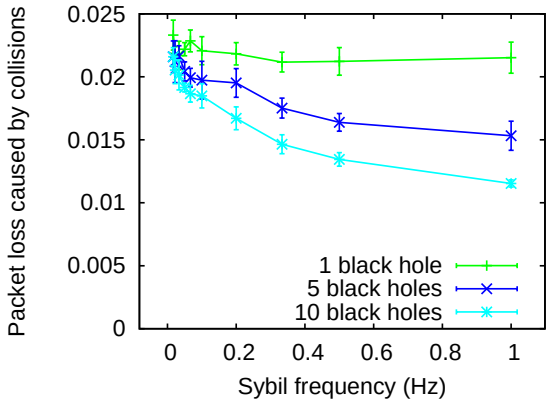
Now, since the probability for packet collisions and route breaks is lower for shorter routes, as shown in Figures 3.6a and 3.8a, the inherent packet loss decreases if the number of misbehaving nodes increases. Also, as shown in Figure 3.7a, the probability for exhausting queuing space decreases if the number of misbehaving nodes increases. This can again be explained by the decreasing route length. If route length decreases, the probability for a route being used by multiple streams in parallel decreases. Thus, the traffic transmitted over a particular route and the queuing space required decreases.

Since the protection of the address-based approach decreases if the Sybil frequency increases, the route length decreases as shown in Figure 3.14b. Thus, for reasons similar to those described above, the inherent packet loss decreases if the Sybil frequency increases. Since the loss caused by one black hole increases only slightly, as shown in Figure 3.4b, the inherent packet loss does not show any significant variations subject to the Sybil frequency in this case. The interpretation is supported by the packet loss caused by collisions on the wireless medium shown in Figure 3.6b, the packet loss caused by exhausted queuing space in the network interface shown in Figure 3.7b, and the packet loss caused by AODV due to route breaks or unavailable destinations shown in Figure 3.8b. All of these metrics show effects comparable to those of an unprotected network described above.

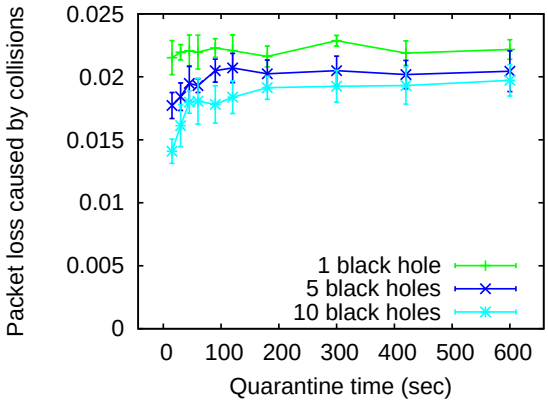
For the location-based intrusion response, the packet loss caused by inherent network properties does not show significant variations subject to the number of misbehaving nodes as well as subject to the quarantine time, as shown in Figure 3.5c. This can be explained by the fact that the packet loss caused by black holes in a network with location-based intrusion response only shows small variations subject to the number of misbehaving nodes as well as subject to the quarantine time as shown in Figure 3.4c. Also the route length for a network with location-based intrusion response varies only slightly, as shown



(a) Defenseless network



(b) Address-based intrusion response

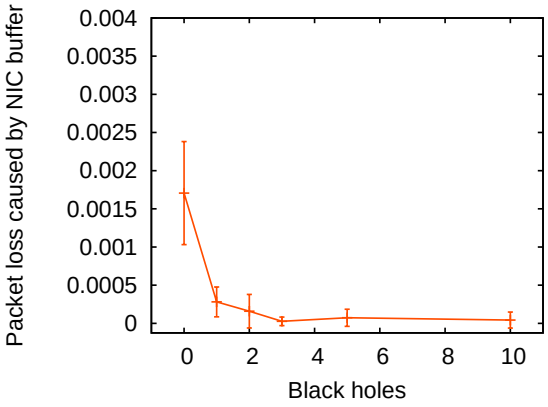


(c) Location-based intrusion response

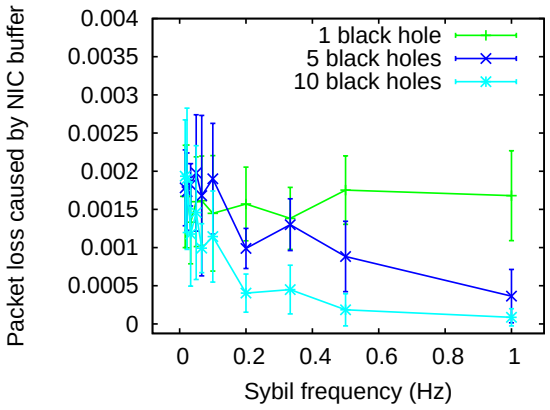
Figure 3.6: Packet loss caused by collisions on the wireless medium

in Figure 3.14c, which supports the interpretation. Only the packet loss caused by collisions on the wireless medium presented in Figure 3.6c shows partially significant variations that correlate to the minor variations in route length.

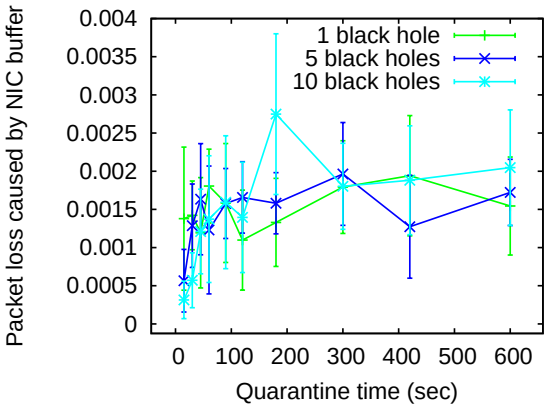
Altogether, the effects we observed for the packet loss caused by inherent network properties, viewed together with the route length, support the interpretation that the address-based intrusion response mechanism can be circumvented by changing addresses.



(a) Defenseless network

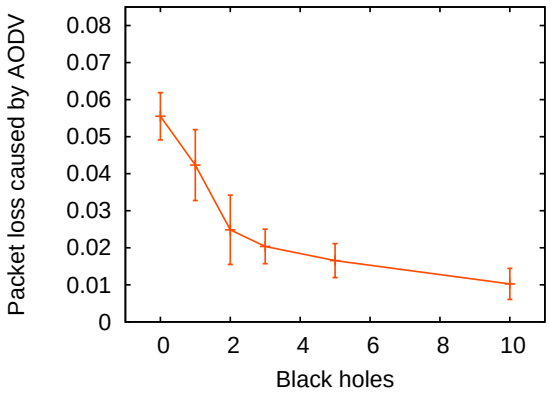


(b) Address-based intrusion response

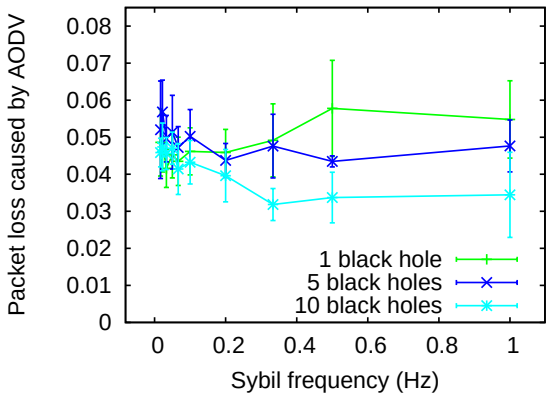


(c) Location-based intrusion response

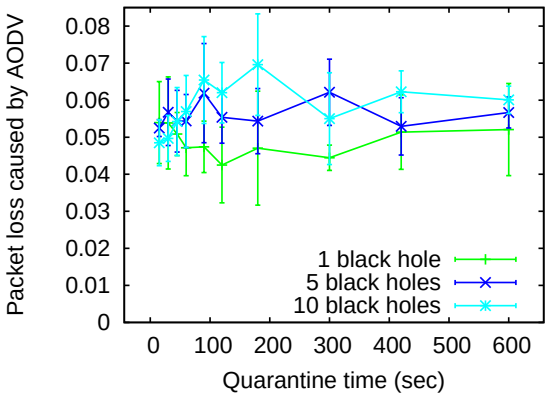
Figure 3.7: Packet loss caused by exhausted queuing space in the network interface



(a) Defenseless network



(b) Address-based intrusion response



(c) Location-based intrusion response

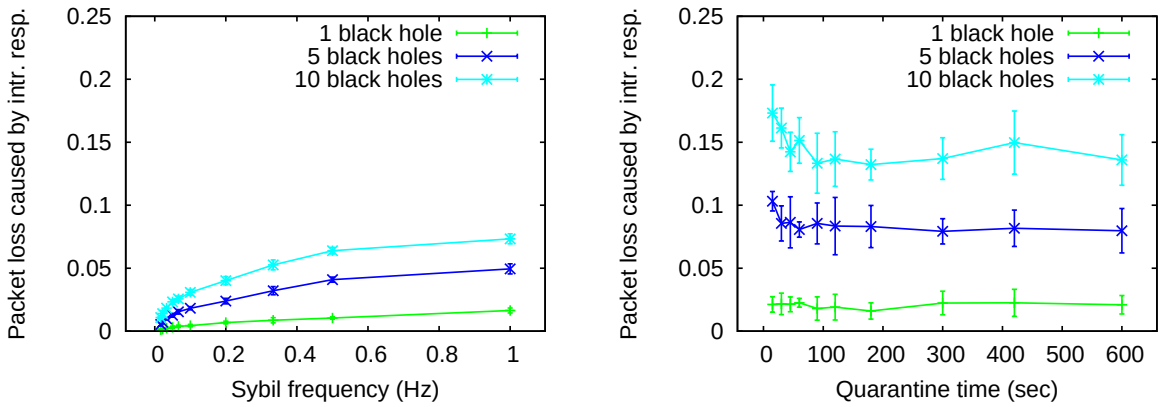
Figure 3.8: Packet loss caused by the AODV routing protocol

As loss caused by the intrusion response system, we consider packets that are not forwarded by a benign node because, otherwise, they would reach a (detected) misbehaving node.

In case of the location-based intrusion response, if a route crosses the boundaries of a quarantined area, the route is interrupted by the first quarantined node. A corresponding route error message is sent to the source. During the time the route error messages takes to reach the source, the source still sends packets using the meanwhile interrupted route. In the first version of location-based intrusion response evaluated here, these packets are dropped by the quarantined node that interrupted the route. Further, if a quarantined node itself has data ready to send, the data is dropped while the node is quarantined. We will discuss a possibility to avoid these reasons for packet loss in Chapter 5. Yet, for now, we charge this loss to the intrusion response system.

In case of the address-based intrusion response, if a route contains a node with an address associated to misbehavior, the node that would have to forward packets to this address interrupts the route. Now, packet loss charged to the address-based intrusion response occurs for reasons similar to those described for location-based intrusion response.

Figure 3.9 shows the resulting packet loss for both alternatives for intrusion response. For the address-based approach, shown in Figure 3.9a, the packet loss caused by intrusion response correlates to the packet loss caused by black holes. The loss increases along with the number of misbehaving nodes and along with the Sybil frequency. Both a higher number of misbehaving nodes and faster changes in addresses of misbehaving nodes lead to an increased number of detections required to exclude misbehaving nodes and, thus, to an increased packet loss caused by intrusion response.



(a) Address-based intrusion response

(b) Location-based intrusion response

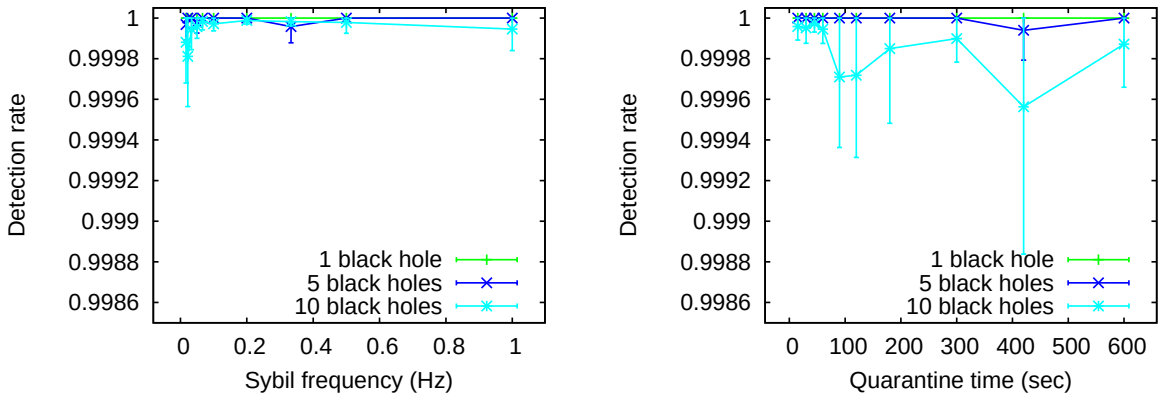
Figure 3.9: Packet loss caused by intrusion response

For the location-based intrusion response, shown in Figure 3.9b, the packet loss caused by intrusion response also correlates to the loss caused by black holes. Yet, as observed for the loss caused by inherent network properties and for similar reasons, the variations are less significant than for the address-based approach. Regarding the number of misbehaving nodes, the packet loss caused by location-based intrusion response increases proportionally from the scenario with one black hole to the scenario with 5 black holes. From 5 to 10 black holes, we observe a reduced growth factor, which can be explained by the fact that the probability for quarantined areas to overlap each other increases along with the number of misbehaving nodes. If quarantined areas overlap, the number of affected benign nodes per quarantined area decreases. Thus, the loss caused if quarantined benign nodes have data ready to send is also reduced.

Comparing the loss caused by the address-based approach and the location-based solution, it stands out that, in particular in scenarios with multiple misbehaving nodes, the address-based solution clearly outperforms the location-based approach. This is due to the fact that for the address-based approach

loss occurs only due to routes that have to be interrupted. For the location-based solution, loss occurs, in addition, if quarantined benign nodes have data ready to send. However, in contrast to the the loss caused by black holes, the loss caused by intrusion response is in our hands. We will present approaches on improving this aspect of the location-based intrusion response in the following chapters.

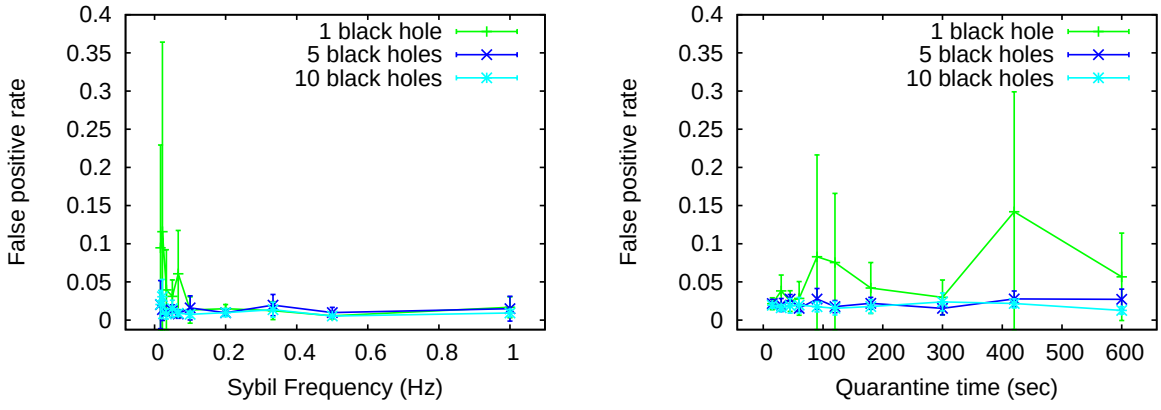
To be sure that equal preconditions regarding intrusion detection apply for both address-based and location-based intrusion response, we monitored the performance of the intrusion detection system in terms of detection rate and false positive rate, as shown in Figures 3.10 and 3.11.



(a) Address-based intrusion response **(b)** Location-based intrusion response

Figure 3.10: Detection rate of the intrusion detection system

Regarding the detection rate, we observe a slightly reduced performance for the location-based approach in a scenario with 10 black holes. Note that this effect is not of statistical significance and the detection rate still is above 99.9%. To explain the small deviation, we have to recall that the detection rate is calculated as correct detections subject to possible detections per monitoring interval. If we now consider that quarantined areas may overlap in scenarios with multiple misbehaving nodes (and, in addition, false positives may occur, further increasing the number of quarantined areas), a case may occur where a misbehaving node is fully quarantined without being detected. In this case, a seperate detection of this misbehaving node is not possible but, also, not required. Still, this rare effect is charged as a failure of intrusion detection.



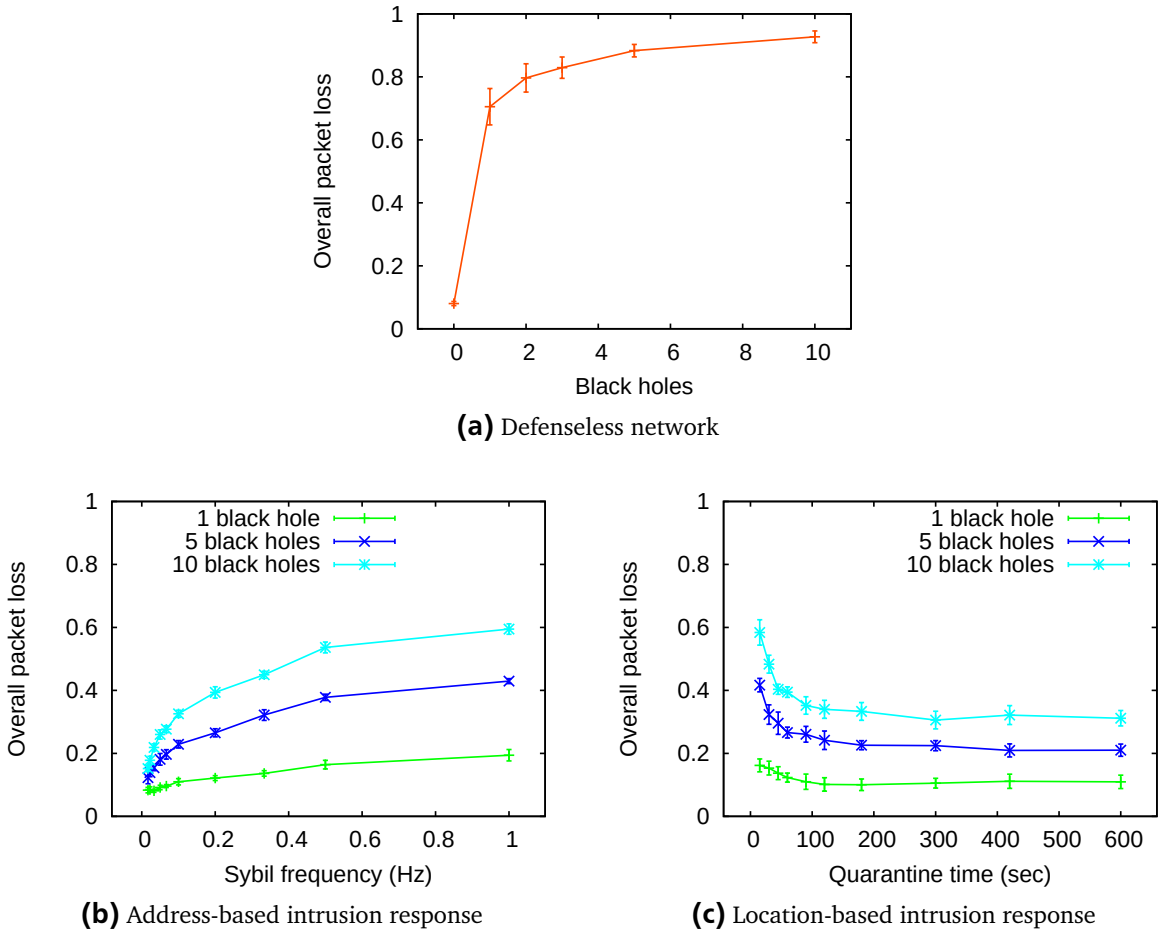
(a) Address-based intrusion response **(b)** Location-based intrusion response

Figure 3.11: False positive rate of the intrusion detection system

Regarding the false positive rate, we observe outliers in scenarios with one misbehaving node for both address-based and location-based intrusion response. Note that this effect is again not of statistical sig-

nificance and the performance of the intrusion detection system remains in the same order of magnitude for both address-based and location-based intrusion response. The effect can be explained by the fact that the inherent packet loss is higher for scenarios with one misbehaving nodes than for scenarios with multiple misbehaving nodes as shown in Figure 3.5. The intrusion detection system does not differentiate between packets that are not forwarded due to misbehavior and packets that are not forwarded due to inherent network conditions. Thus, a high loss due to inherent network conditions may lead to false positive detections.

Overall Packet Loss and Delivery Ratios



(a) Defenseless network

(b) Address-based intrusion response

(c) Location-based intrusion response

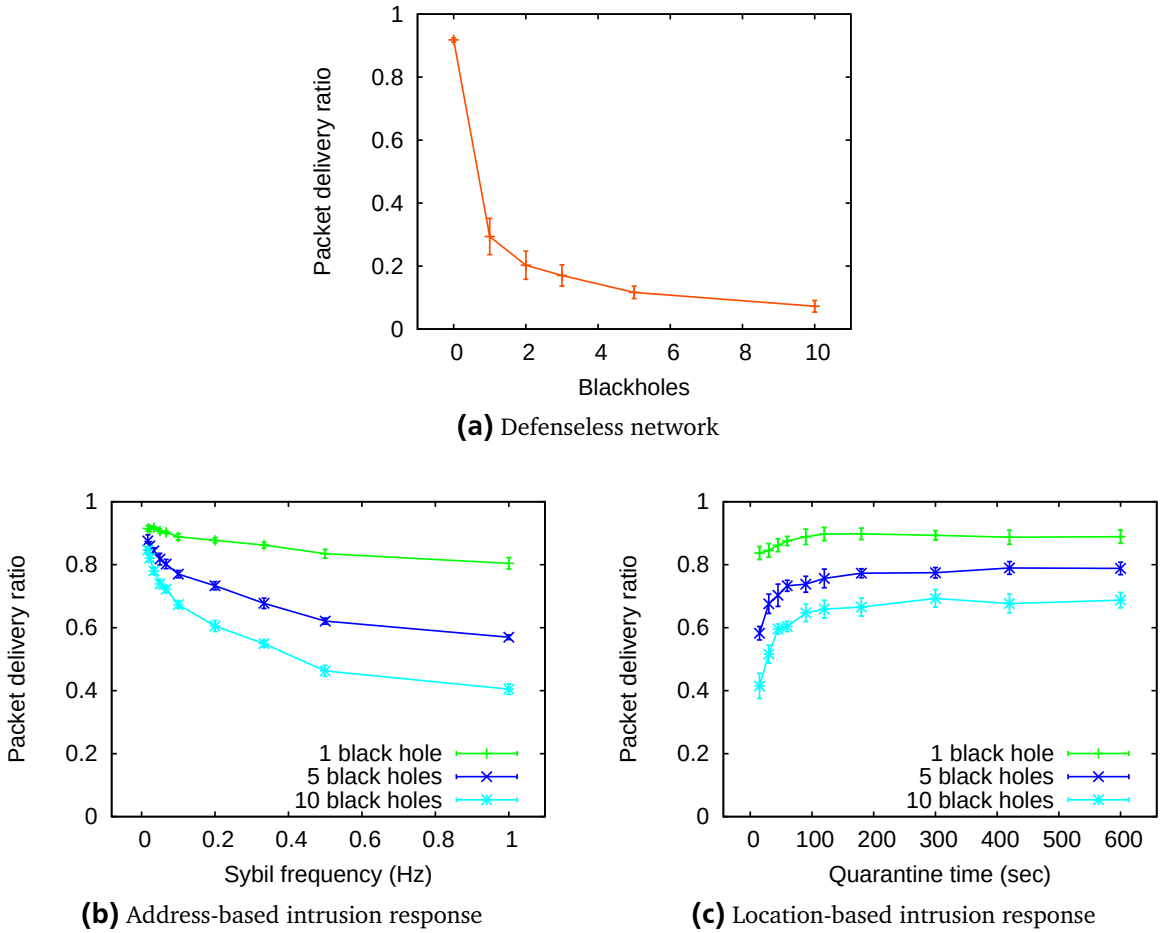
Figure 3.12: Overall packet loss

The overall packet loss as sum of (1) the loss caused by black holes, (2) the packet loss caused by inherent network properties, and (3) the packet loss caused by intrusion response is shown in Figure 3.12. To interpret the results, we have to recall that the Sybil frequency is part of the misbehavior and can be optimized by a potential attacker in order to maximize the impact of the attack, thus, in the particular case of a black hole attack, to maximize packet loss in the network. The quarantine time, on the other hand, is part of the intrusion response system and can be optimized to minimize the impact of misbehavior. Comparing the corresponding results, we see that the location-based intrusion response clearly outperforms the address-based solution. As an example, for a scenario with 10 misbehaving nodes, we observe an overall packet loss of more than 90% in a defenseless network as shown in Figure 3.12a. If we consider a maximized attack performance for a Sybil frequency of one address change per second, the overall loss can be reduced to about 60% with the address-based intrusion response.

For an optimal choice of the quarantine time, the overall loss can be reduced to about 30% for the location-based approach.

We further have to remember that a considerable part of the overall loss observed for the location-based intrusion response is the loss caused by intrusion response itself which is still under control of the intrusion response system. To improve the location-based approach we, thus, focus on reducing this loss in Chapters 4 and 5.

To verify the correctness of the overall loss, we present the delivery ratio observed in the network in Figure 3.13. Note that, in the simulation tool used for the evaluation, packet delivery is monitored separately, independent from packet loss. Failures in the measurement of the packet loss would, thus, be recognizable in the delivery ratio. Since the delivery ratios match the overall packet loss, we assume correctness of the implemented monitoring architecture as well as completeness of the loss metrics considered.



(a) Defenseless network

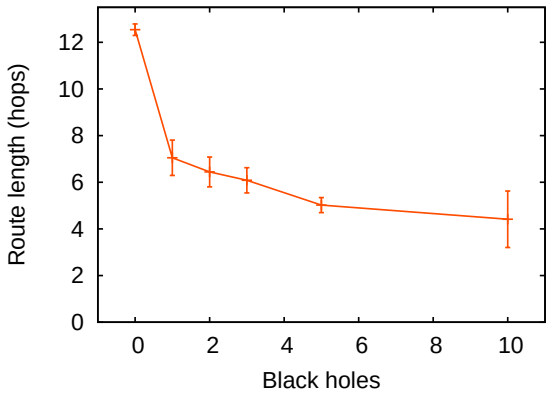
(b) Address-based intrusion response

(c) Location-based intrusion response

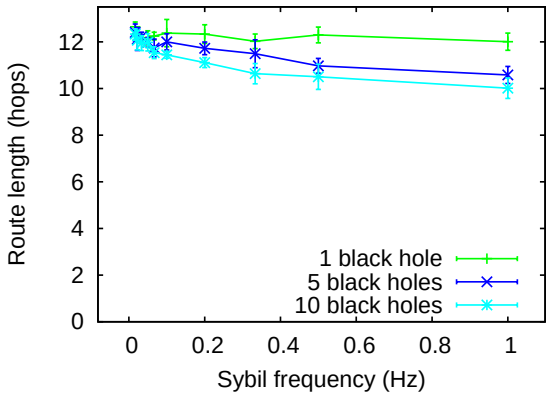
Figure 3.13: Overall delivery ratio

Route Lengths

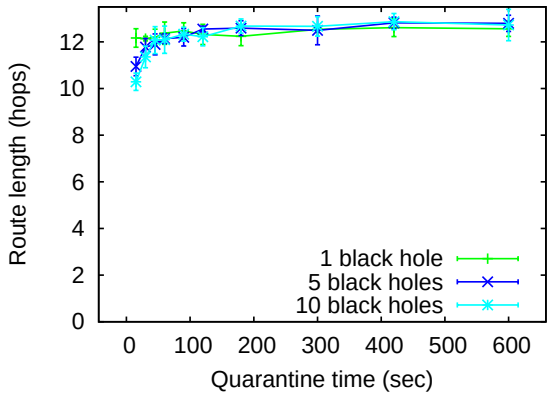
The length of routes over which packets can be transmitted successfully, without being affected by misbehaving nodes, can be directly used to identify whether and to what extent a network is partitioned. Considering a defenseless network, shown in Figure 3.14a, we see that the route length decreases significantly if we increase the number of misbehaving nodes. In a network without misbehaving nodes, we observe an average route length of about 13 hops, which decreases to about 4 hops for scenarios with 10 misbehaving nodes. Thus, if we introduce black holes, the network becomes partitioned logically and communication is possible only if source and destination are in close proximity. This effect can be explained by the expanding ring search described in Section 2.2.3, which is used by the AODV routing protocol to limit the overhead introduced by the broadcast-based route discovery process. During the expanding ring search, the broadcast is, at the beginning, limited to the direct neighborhood of the source node and extended to the whole network only if the destination can not be found nearby the source. Thus, if the destination node is located in a ring closer to the source than the ring that contains the next misbehaving node, the misbehaving node does not receive the corresponding route request and is not able to tamper with the route. The further away the destination is located from the source, the higher the probability is for a black hole to receive the route request. We further quantify this effect in Section 3.3.



(a) Defenseless network



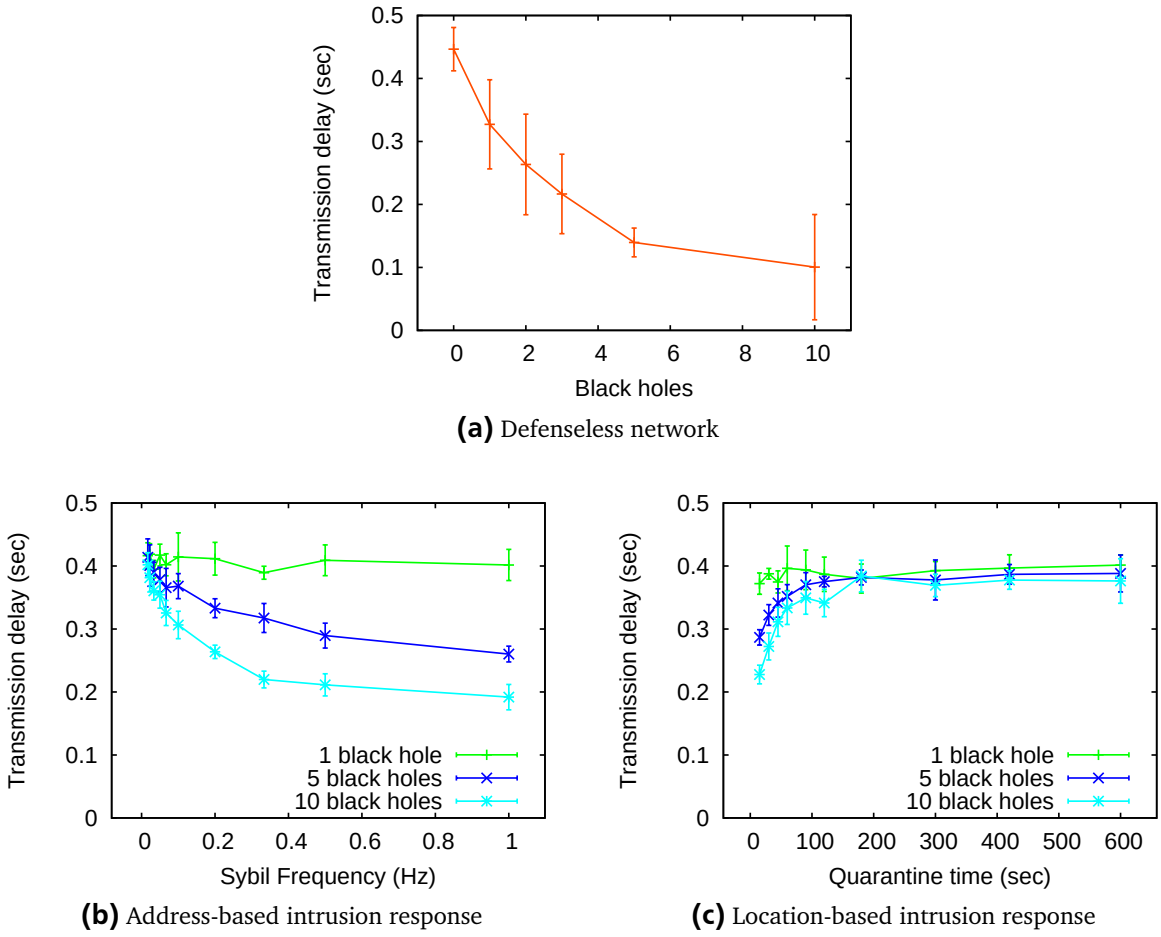
(b) Address-based intrusion response



(c) Location-based intrusion response

Figure 3.14: Route length

Figure 3.15 shows the transmission delay of packets transmitted successfully, measured from sending application to receiving application. The transmission delay consists of the routing delay on the source shown in Figure 3.16, and the propagation delay of the network shown in Figure 3.17. The routing delay on the source node is the time it takes the AODV routing protocol to establish a route to the destination. Note that the routing delay does not affect each packet. Only packets sent at the beginning of a stream have to be buffered until a route is discovered. However, we present the average routing delay experienced by all packets. The propagation delay denotes the time a packet takes to travel through the network to the receiving application.



(a) Defenseless network

(b) Address-based intrusion response

(c) Location-based intrusion response

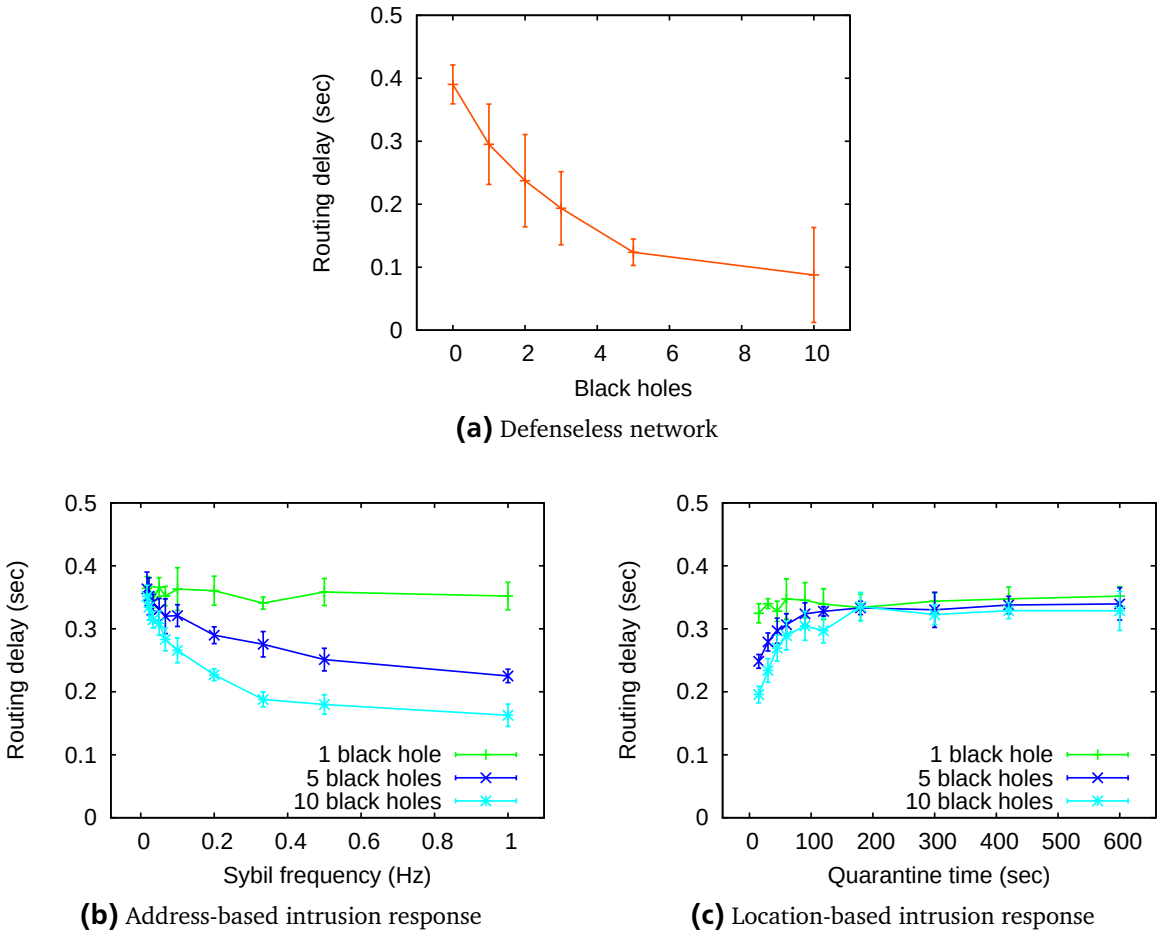
Figure 3.15: Transmission delay from application to application

For obvious reasons, the routing and propagation delays and, thus, the overall transmission delay directly depend on the route length in hops and do not show unexpected results. It only stands out that the maximum delay for both the address-based and the location-based intrusion response are slightly lower than for a defenseless network without malicious nodes.

For the routing delay, this can be explained by the fact that, in case a misbehaving node was part of a route and was detected, the expanding ring search of AODV does not start from the very beginning, but from the ring that contained the misbehaving node. Thus, the routing delay for packets that are transmitted successfully after misbehavior was detected decreases.

For the propagation delay, this can be explained by the fact that the delays are measured on a per packet basis. Thus, if less packets are transmitted successfully via long routes, the average delay decreases. In scenarios with misbehaving nodes, long routes are affected by misbehaving nodes with a

higher probability than short routes, which is due to the expanding ring search mechanism of AODV. Thus, on long routes more packets get lost due to intrusion response than on short routes. Therefore, although the length of routes established successfully in scenarios with misbehavior and optimal intrusion response matches the route length in a network without misbehavior and intrusion response, the average delay is slightly lower.

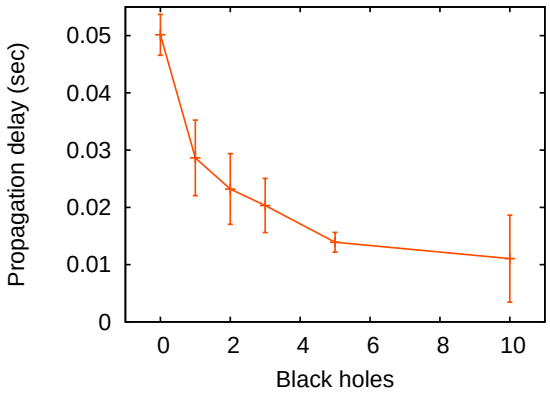


(a) Defenseless network

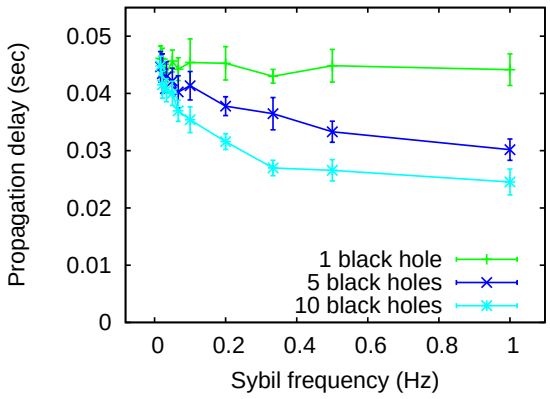
(b) Address-based intrusion response

(c) Location-based intrusion response

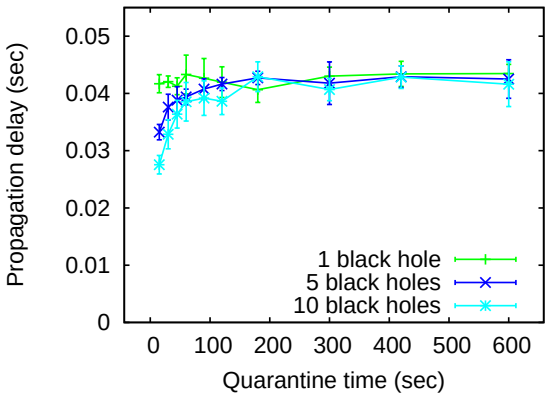
Figure 3.16: Routing delay on source node



(a) Defenseless network



(b) Address-based intrusion response



(c) Location-based intrusion response

Figure 3.17: Propagation delay from source to destination

3.3 Analytical Validation

In this section, to double-check the correctness of the simulation results, we develop an analytical model describing the effects of the black hole attack and of the location-based intrusion response on the packet loss observed in the network. After defining basic assumptions, we start with modeling the expanding ring search behavior of the AODV routing protocol which leads to the probability for a black hole being included in a route. Based on this, we describe the packet loss caused by black holes in a defenseless network without intrusion detection and intrusion response. We continue with modeling the packet loss caused by black holes with activated location-based intrusion response. For this, we consider the time the intrusion detection system requires to detect a black hole and the mobility of nodes. A model for the packet loss caused by the location-based intrusion response itself concludes this section.

Table 3.4: Notations of formulae

t_{mon}	Monitoring interval of the intrusion detection system in seconds
$n_{X,rec}$	Number of packets Node X received during t_{mon}
$n_{X,forw}$	Number of packets Node X forwarded during t_{mon}
w_{bal}	Factor to balance $n_{X,rec}$ and $n_{X,forw}$
$R_{X,Y}$	Rating for Node Y determined by the intrusion detection system of Node X
$thres_{black}$	Threshold of R_Y for classification as black hole
r_{quar}	Radius of quarantined areas
t_{detect}	Time needed to detect a black hole
t_{reset}	Time after which quarantined areas are revoked
r_{trans}	Transmission range of nodes
A_{net}	Size of the network
l	Side length of the network area
d_{hop}	Average distance per hop
n_{total}	Total number of nodes in the network
n_{black}	Number of black hole nodes in the network
n_s	Number of nodes reached in step s of the ring search
ρ	Network density in nodes per area

3.3.1 Assumptions

In the following, we assume

- a square network area A_{net} with side length l,
- a circular transmission area with radius r_{trans},
- quarantined areas to match the transmission area of the corresponding misbehaving node at the time the quarantined area is established,
- a geometrically uniformly distributed random placement of benign nodes and black holes within the network area,
- randomly selected sources and destinations of traffic,
- a random distribution of traffic patterns among all nodes, that is, the network load is constant and nodes can not be distinguished by their communication,

- a connected network, that is, a route between any two nodes can be established at any time.

3.3.2 Expanding Ring Search

For modeling the packet loss, we need to describe the expanding ring search of AODV in terms of the number of new nodes a route request message reaches in each step as shown in Figure 3.18. For this, let A_h be the circular area covering all nodes located in a distance of at most h hops to the source of the route request, where $A_0 = 0$ and $\Delta A_h = A_h - A_{h-1}$. Depending on the actual network configuration, a fix average geometric per-hop routing progress $0 < d_{hop} < r_{trans}$ can be assumed. For our model, we use

$$d_{hop} = \frac{r_{trans}}{\sqrt{2}}$$

which was obtained experimentally in [36] for a realistic mobile ad hoc network configuration. This leads to a first estimate $\Delta\Lambda_h$ of ΔA_h defined as

$$\begin{aligned} \Delta\Lambda_h &= \Lambda_h - \Lambda_{h-1} = \pi(h \cdot d_{hop})^2 - \pi((h-1)d_{hop})^2 \\ &= (2h-1)\pi \cdot d_{hop}^2 \end{aligned}$$

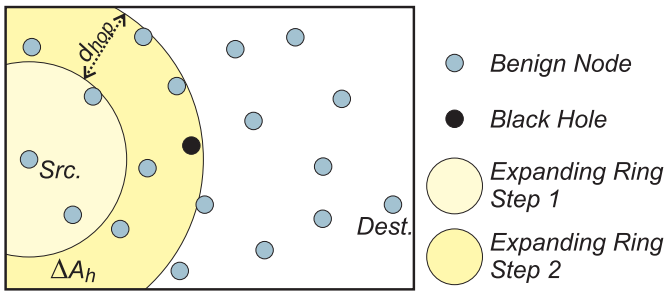


Figure 3.18: Schematic representation of the area covered by the expanding ring search subject to d_{hop}. Note that, for reasons of presentation, dimensions do not match the assumptions made for the model.

This description holds if the ring search does not leave the simulation area. To include edge effects for nodes being located at the borders of the simulation area, we use a four-step approximation. For a ring radius $r = h \cdot d_{hop} < 0.5 \cdot l$ of less than half the side length l of the simulation area, the ring can be contained fully in the simulation area in the best case, for a node being located at the center. In the worst case, for a node being located at a corner, only one quarter of the ring reaches into the simulation area. For the approximation, we assume that both cases and any in between are equally likely. Thus, for $r = h \cdot d_{hop} < 0.5 \cdot l$ we obtain

$$\Delta A_h = \frac{1}{2}\left(\Lambda_h + \frac{1}{4}\Lambda_h\right)$$

In analogy, we approximate the cases for a ring radius between half side length and side length, between side length and diagonal length, and larger than diagonal length. Altogether, we obtain

$$\Delta A_h = \begin{cases} \frac{1}{2}\left(\Lambda_h + \frac{1}{4}\Lambda_h\right) & \text{if } h \cdot d_{hop} \leq \frac{l}{2} \\ \frac{1}{8}\Lambda_h & \text{if } \frac{l}{2} < h \cdot d_{hop} \leq l \\ \frac{1}{16}\Lambda_h & \text{if } l < h \cdot d_{hop} \leq \sqrt{2l^2} \\ 0 & \text{otherwise} \end{cases}$$

As described in Section 2.2.3, the expanding ring search is performed in consecutive steps in each of which the time to live of a route request is incremented. Let h_s be the time to live in hops for step s. With the default parameters specified in [78], h_s is given by the following table (we set $h_0 = 0$ for reasons of simplification).

s	0	1	2	3	4	5
h_s	0	1	3	5	7	35

For modeling the number n_s of nodes that are reached in step s of the expanding ring search, let n_{total} be the total number of nodes and A_{net} be the total area of the network. Because nodes are assumed to be distributed uniformly, the geometric density of nodes is given by

$$\rho = \frac{n_{total}}{A_{net}}$$

We get

$$\begin{aligned} n_0 &= 0 \\ n_s &= n_{s-1} + \Delta n_s \text{ where } \Delta n_s = \sum_{i=h_{s-1}+1}^{h_s} \Delta A_i \cdot \rho \end{aligned}$$

3.3.3 Packet Loss Caused by Black Holes in a Defenseless Network

To model the packet loss caused by black holes, we start with describing the probability $p_{black}(s)$ for a route request reaching at least one black hole until step s of the expanding ring search. If we assume n_{black} black holes in n_{total} total nodes, we have $n_{total} - n_{black}$ benign nodes in the network. Thus, the number of possible combinations to select the n_s nodes that are reached in step s of the ring search only from benign nodes is

$$x = \binom{n_{total} - n_{black}}{n_s}$$

The number of combinations for selecting the n_s nodes from n_{total} total nodes is

$$y = \binom{n_{total}}{n_s}$$

Now, $x \cdot y^{-1}$ describes the probability that a route request reaches only benign nodes until step s of the ring search. For the probability $p_{black}(s)$ that a route request reaches at least one black hole until step s of the ring search, we have to consider the special case of $n_s > n_{total} - n_{black}$, that is, the number of nodes reached in step s exceeds the number of benign nodes. Certainly, the route request reaches a black hole in this case. Altogether, we get

$$p_{black}(s) = \begin{cases} 1 & \text{if } n_s > n_{total} - n_{black} \\ 1 - \frac{\binom{n_{total} - n_{black}}{n_s}}{\binom{n_{total}}{n_s}} & \text{otherwise} \end{cases}$$

We assume a random selection of traffic patterns as well as of source and destination nodes. Hence, the probability for a data packet to get lost due to a black hole equals the probability for a black hole being part of the corresponding route between source and destination. To determine this, we need to calculate the probability $p_{dest}(s)$ for reaching the destination node in step s of the ring search. Since destinations are assumed to be chosen randomly, this probability correlates to Δn_s. We obtain

$$p_{dest}(s) = \frac{\Delta n_s}{n_{total}}$$

With this, the probability for the route request reaching the destination in step s and at least one black hole until step s of the ring search is given by $p_{dest}(s) \cdot p_{black}(s)$. For the overall probability $p_{loss,black}$ for data packets to get lost due to a black hole, we need to consider all steps of the expanding ring search. We get

$$p_{loss,black} = \sum_{i=1}^{5} p_{dest}(i) \cdot p_{black}(i)$$

3.3.4 Packet Loss Caused by Black Holes with Intrusion Detection and Intrusion Response

Although the location-based intrusion response excludes black hole nodes from the network, the influence of black holes can only be mitigated and not be thwarted completely. Thus, the probability $p_{loss,defended}$ for packet loss caused by black holes despite active security measures can be described as

$$p_{loss,defended} = p_{loss,black} \cdot p_{IRSfail}$$

Here, $p_{IRSfail}$ denotes the probability that the intrusion response system fails to prevent a black hole from dropping packets. As main factors on $p_{IRSfail}$ we identified the detection time of the intrusion detection system and the mobility of nodes. Both lead to independent parts $p_{IRSfail,detect}$ and $p_{IRSfail,move}$ of $p_{IRSfail}$. Thus, we obtain

$$p_{IRSfail} = p_{IRSfail,detect} + p_{IRSfail,move}$$

Impact of the Intrusion Detection System

To detect ongoing misbehavior, the intrusion detection system has to monitor the suspicious node for a certain time. In our scenario, the black hole may continue dropping packets during this time. To describe the resulting probability $p_{IRSfail,detect}$ (as a part of $p_{IRSfail}$) for packet loss during the time, the intrusion detection system needs to identify misbehavior, we start by modeling the detection time. For our intrusion detection system as introduced in Section 3.1.4, the detection time is given as

$$t_{detect} = n_{mon} \cdot t_{mon}$$

where n_{mon} denotes the number of monitoring intervals required until a black hole is detected and t_{mon} denotes the duration of a monitoring interval.

As described, a node Y is classified as a black hole if the rating $r_{X,Y}$ of Y by Node X exceeds the threshold $thres_{black}$. Since, in our case, a black hole node does not forward any packets, $r_{X,Y}$ is defined as

$$r_{X,Y} = \frac{n_{X,rec}}{w_{bal}}$$

per monitoring interval t_{mon}. If we consider a steadily loaded network, the traffic during the detection time is (nearly) constant at a rate λ. Thus, $n_{X,rec} = \lambda \cdot t_{mon}$. If a black hole is detected,

$$n_{mon} \cdot \frac{\lambda \cdot t_{mon}}{w_{bal}} > thres_{black}$$

holds. The number n_{mon} of monitoring intervals required to detect a black hole is then defined as

$$n_{mon} = \frac{thres_{black} \cdot w_{bal}}{\lambda \cdot t_{mon}}$$

Thus, for the detection time of the intrusion detection system, we obtain

$$t_{detect} = \frac{thres_{black} \cdot w_{bal}}{\lambda}$$

After a black hole is detected, the corresponding quarantined area excludes the black hole from the network for the time t_{reset}. Altogether, a black hole can be active for the time t_{detect} as part of the total detection-protection-period $t_{detect} + t_{reset}$. Thus, for the probability $p_{IRSfail,detect}$ of losing packets due to n_{black} black holes during the detection time of the intrusion detection system, we obtain

$$p_{IRSfail,detect} = n_{black} \cdot \frac{t_{detect}}{t_{detect} + t_{reset}}$$

Impact of Node Mobility

We assume that quarantined areas can not be adapted directly when a node moves since tracking is not possible when a node is quarantined. Thus, as shown in Figure 3.19, mobility of a black hole leads to a newly affected area A_{affect}. Nodes in this area are not aware of the black hole and will forward route request messages without restrictions.

The probability $p_{IRSfail,move}$ for packet loss due to node mobility can be modeled based on the number n_{affect} of nodes located in A_{affect}. Each of these nodes has to perform a detection of the black hole, leading to a corresponding multiple of $p_{IRSfail,detect}$ as described in the previous section.

The angle α, as shown in Figure 3.19, can be determined by

$$\alpha = 2 \cdot \arccos\left(\frac{d}{2r_{trans}}\right)$$

where d denotes the distance between the location where the black hole was first detected and its new location. With this, we can determine the area A_{seg} of the circular segment defined by the quarantine zone and A_{affect}. We get

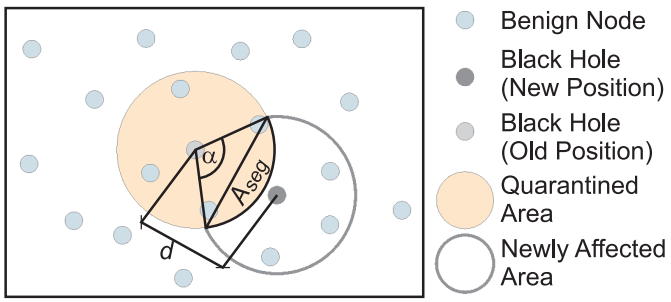


Figure 3.19: Influence of black hole mobility on the IRS

$$A_{seg} = \frac{r_{trans}^2}{2} \cdot (\alpha - \sin(\alpha))$$

thus,

$$A_{affect} = \pi r_{trans}^2 - 2 \cdot A_{seg}$$

Since nodes are distributed uniformly within the network area, we get

$$n_{affect} = A_{affect} \cdot \rho$$

Thus, the loss due to node mobility can be described as

$$p_{IRSfail,move} = n_{affect} \cdot n_{black} \cdot p_{IRSfail,detect}$$

3.3.5 Packet Loss Caused by the Intrusion Response System

Packet loss caused by the intrusion response system arises from benign nodes being source or destination of a communication while located in quarantined areas. To model the probability $p_{loss,IRS}$ for packet loss due to intrusion response, we, thus, have to determine the number of nodes that are located in quarantined areas.

To calculate the total area $A_{quar}(n_{black})$ that is covered by quarantined areas, we use a simple heuristic for the probability $p_{lap}(n_{black})$ that quarantined areas overlap. This is the case if the distance between two black holes is less than $2r_{trans}$. We define

$$\begin{aligned} p_{lap}(1) &= 0 \\ p_{lap}(i) &= \frac{A_{lap}(i-1)}{A_{net}} \text{ if } i > 1 \end{aligned}$$

The areas $A_{lap}(n_{black})$ that are needed to determine $p_{lap}(n_{black})$ are defined as

$$\begin{aligned} A_{lap}(1) &= \pi \cdot (2 \cdot r_{trans})^2 \\ A_{lap}(i) &= A_{lap}(i-1) + (1 - p_{lap}(i)) \cdot \pi (2 \cdot r_{trans})^2 \text{ if } i > 1 \end{aligned}$$

With this, $A_{quar}(n_{black})$ can be calculated as

$$\begin{aligned} A_{quar}(1) &= \pi \cdot r_{trans} \\ A_{quar}(i) &= A_{quar}(i-1) + \pi r_{trans}^2 \cdot (1 - p_{lap}(i)) \text{ if } i > 1 \end{aligned}$$

To determine $p_{loss,IRS}$, we now have to describe the probability of either source, or destination, or both nodes being located within a quarantined area. Since nodes are distributed uniformly, the probability that an arbitrarily chosen node is available (that is, not quarantined) is given as

$$p_{avail} = 1 - \frac{A_{quar}(n_{black})}{A_{net}}$$

The probability that both source and destination are available is p_{avail}^2. Thus, we obtain

$$p_{loss,IRS} = 1 - p_{avail}^2$$

3.3.6 Comparison of Model Predictions and Simulation Results

To validate our work, we compare predictions made by the model with the results of the simulation study we presented in 3.2. To visually demonstrate the accuracy achieved, we show the model predictions as curves together with the 95% confidence bars taken from the simulation results in Figure 3.20. To repeat the details required for the model, the simulation results were obtained from a scenario consisting of $n_{total} = 1000$ nodes with $n_{black} \in \{1, 2, 3, 5, 10\}$ black holes on a square simulation area A_{net} with $l = 4750m$ side length. The nodes moved continuously according to a random waypoint mobility model with a speed between 1 and 2 meters per second. This leads to an average distance

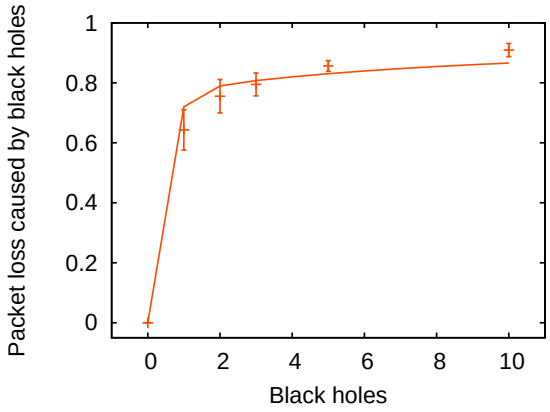
$$d = \frac{t_{reset} \cdot \text{meters}}{6 \cdot \text{second}}$$

from the location at which a black hole was detected to its new location when the quarantined area is revoked. We used a constant bitrate traffic pattern with a network load of 20 streams in parallel consisting of packets with a size of 512 byte at a rate of 2048 bytes per second. With a detection threshold $thres_{black} = 10$ determined as an optimum in preliminary simulations for the scenario given, the resulting detection time of the intrusion detection system is $t_{detect} = 1s$. The reset intervals for quarantined areas were taken from $t_{reset} \in \{15, 30, 45, 60, 90, 120, 180, 300, 420, 600\}$ in seconds.

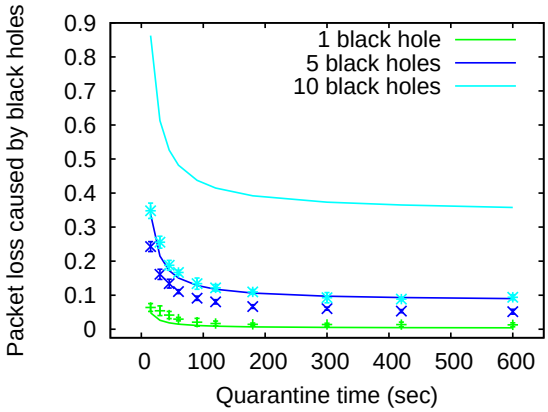
The comparison of the loss caused by black holes in a defenseless network is shown in Figure 3.20a. For the 2, 3, and 5 black hole scenarios, the model prediction is within the confidence interval of the simulation results. The small deviation for the setups with 1 and 10 black holes can be explained by the inaccuracy of the heuristic we used to model edge effects of the expanding ring search. Yet, the error of the model prediction is alway less than two times the simulation confidence as shown in Table 3.5.

Figure 3.20b shows the comparison of the loss caused by black holes in a network with intrusion detection and location-based intrusion response. Please note that, for reasons of readability, only the results for the 1, 5, and 10 black hole scenarios are presented.

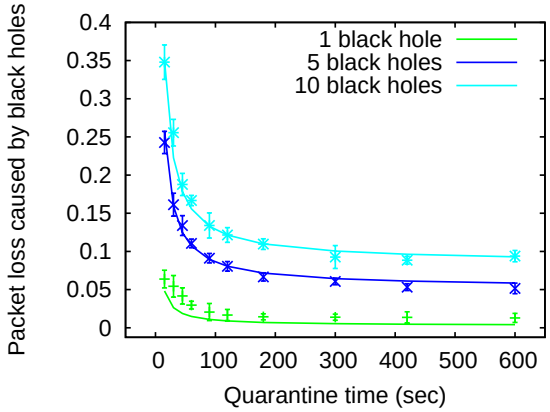
While prediction and confidence intervals (that is, simulation results) match well for the 1 black hole scenario (and, without presentation, for the 2 and 3 black hole scenarios), it stands out that, for the 5 and 10 black hole setups, the prediction and the simulation results differ strongly (note that the curve matching the 10 black hole confidence bars belongs to the 5 black hole prediction). Rerunning and



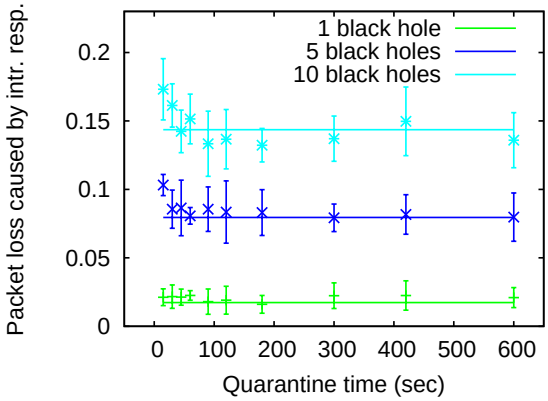
(a) Loss due to black holes without IDS/IRS



(b) Loss due to black holes with IDS/IRS



(c) (b) adapted to black hole activity



(d) Loss due to IRS

Figure 3.20: Comparison of model predictions (curves) and simulation results (confidence bars)

Table 3.5: Comparison of model predictions and simulation results for the packet loss due to black holes in a defenseless network

n_{black}	model prediction x	simulation result y	simulation confidence z	$x - y$	$\frac{x-y}{z}$
1	0.720	0.643	0.067	0.077	1.144
2	0.789	0.755	0.056	0.034	0.611
3	0.807	0.795	0.038	0.013	0.334
5	0.830	0.856	0.018	-0.026	-1.454
10	0.866	0.909	0.022	-0.043	-1.964

tracing the simulations, we found that in the scenario with 5 black holes only 4 and in the scenario with 10 black holes only 5 are active at the same time. We put this down to overlapping quarantined areas and to effects of black holes located at the edges of the simulation area. If we instantiate the model accordingly (4 / 10 black holes in the model for the 5 / 10 black hole simulation setups), we obtain a match for all scenarios as shown in Figure 3.20c.

The comparison of the loss that is caused by the intrusion response system itself is shown in Figure 3.20d. Again, we obtain a reasonable match of model and simulation. Yet, in this case the simulation results show significant variances, which are due to false positives of the intrusion detection system that are not considered by the model.

3.4 Conclusion - Location-based Intrusion Response

In this chapter, we showed that intrusion response in mobile ad hoc networks should not be performed based on network addresses of misbehaving nodes. By means of simulation, we proved that an address-based intrusion response can be subverted easily and effectively by a Sybil attack. To overcome this drawback, we proposed a location-based intrusion response system. Our studies show that this intrusion response strategy is not affected by Sybil attacks and manages to maintain the basic functionality of the network even in the presence of multiple malicious nodes. Note that our results hold for the assumptions we made for the evaluation and the analytical validation in Sections 3.2 and 3.3. In particular, we considered a connected network and a low number of misbehaving nodes. In more adverse conditions, intrusion response might be insufficient to recover network functionality. In these scenarios, robust, broadcast-based routing protocols such as [81] or disruption-tolerant, mobility-based message dissemination as proposed in [33, 109] can be deployed. Still, supporting these protocols by quarantined areas to, for example, stop the spreading of flooding attacks can be reasonable.

Although the location-based intrusion response is insusceptible to changes in addresses of misbehaving nodes, it is limited by inherent drawbacks. In its naïve version, quarantined areas are of the size of the transmission/reception range of nodes. As a first drawback, this results in a considerable number of benign nodes that are located in quarantined areas and, thus, are excluded from the network. Second, routes between benign nodes outside quarantined areas are interrupted due to intermediate nodes being quarantined. We consider these drawbacks, as shown schematically in Figure 3.21, to be the most severe ones. Therefore, in the following chapters, we will be concerned with how to optimize the performance of the location-based intrusion response mechanism.

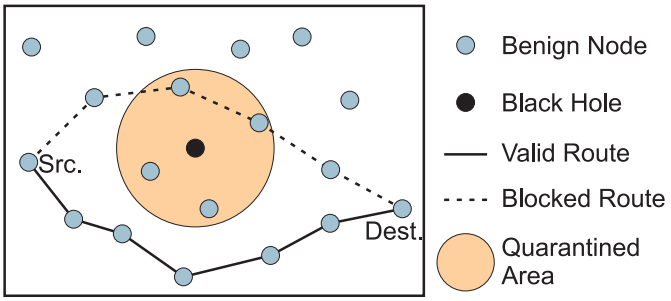


Figure 3.21: Drawbacks of the location-based IRS

At this point, a discussion on the feasibility of mobile ad hoc networks when faced with drop rates around 10% even in the best case may arise. Opinions about this differ widely. Our goal is to offer reliable (basic) communication and information services in scenarios where a communication infrastructure is not available, as it may be the case in large-scale disaster scenarios. Here, mobile ad hoc networks can be deployed to enable communication at all. When taking a look behind the scenes, this could be a useful innovation even if performance (bandwidth) is 'wasted' due to heavy error correction mechanisms. We also remember similar discussions in the early times of (infrastructure based) IEEE 802.11 wireless networks. In this case, mechanisms that are able to deal with (most of) the drawbacks have been successfully developed such that these networks today offer a valuable alternative for wired networks in many scenarios.

To be able to predict the behavior of the location-based intrusion response for other parameter sets than those studied in this chapter, we developed an analytical model that describes the effects of black hole attacks and of the location-based intrusion response. Based on a combined geometric and stochastic approach, we developed an analytical model for the routing process as well as for the packet loss caused by the misbehavior and by the countermeasures. The comparison of model predictions and simulation results shows that the model developed produces reasonable predictions despite following an elementary approach.

4 Supporting Location-based Intrusion Response in Mobile Ad Hoc Networks with Adaptive Transmission Power

In this chapter, we analyze how an adaptive transmission power of benign devices can be used to reduce the size of quarantined areas in order to mitigate the drawbacks of the location-based intrusion response mechanism for mobile ad hoc networks discussed in Section 3.4. In the following, we introduce different approaches for extending the location-based intrusion response by an adaptive transmission power. We evaluate the approaches presented with focus on discovering positive and negative effects of an adaptive transmission power with respect to the size of quarantined areas and node mobility.

4.1 Architecture

In this section, we describe the add-ons for adaptive transmission power to the first, naïve version of location-based intrusion response introduced in Chapter 3. We provide mathematical descriptions of the adaptive transmission power based on the free space model for wireless signal propagation explained in Section 2.2.1.

4.1.1 Naïve Location-based Intrusion Response

In the first, naïve version of of the location-based intrusion response mechanism as introduced in Chapter 3, transmission power is static. Thus, the radius r_{quar} of a quarantined area has to cover at least the distance after which a signal can not be received correctly anymore. For the naïve approach, r_{quar} is calculated as

$$r_{quar} = \sqrt{P_s \cdot G_s \cdot G_r \cdot \frac{\lambda^2}{4 \cdot \pi^2 \cdot P_{min}}}$$

With the settings used for the evaluation of the naïve approach as listed in Table 3.1, we obtain $r_{quar} = 250m$. To offer a security buffer for node mobility, r_{quar} can be increased appropriately which will be subject of our evaluation.

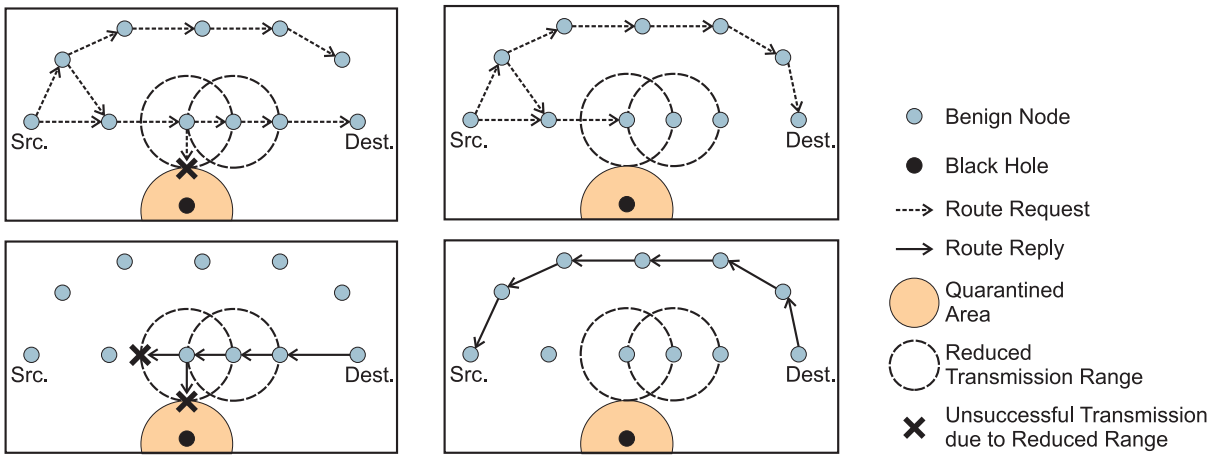
4.1.2 Location-based Intrusion Response with Adaptive Transmission Power

If benign nodes are capable of adapting their transmission power P_S, we may reduce the size of quarantined areas and, thus, minimize the number of benign nodes that are excluded from the network. In theory, it is possible to reduce transmission power such that a quarantined area can be reduced to one single point in Cartesian space. Still, it is reasonable to have a security buffer for node mobility. Being given r_{quar} and the distance d_{center} from a node's location to the center of the closest quarantined area, we can calculate P_S as

$$P_S(d_{center}) < \frac{P_{min}}{G_s \cdot G_r} \cdot \frac{4 \cdot \pi^2 \cdot (d_{center} - r_{quar})^2}{\lambda^2}$$

4.1.3 Location-based Intrusion Response with Adaptive Transmission Power and Asymmetry Prevention

Although we assume an IEEE 802.11 medium access control protocol based on the request to send/clear to send handshake outlined in Section 2.2.2, we introduce asymmetric links by reducing the transmission power. While route request messages are sent as a broadcast without the handshake sequence that would prevent asymmetric links, route reply messages are sent as unicast messages. This may lead to scenarios where a node that operates with reduced transmission power is able to receive and forward a route request message, but will not be able to forward the corresponding route reply message on its way back to the initiator of the route request. To prevent this, we further extend the adaptive transmission power such that a node operating at reduced transmission power only forwards a route request message if the distance to the node from which it received the route request is less than d_{center}. Figure 4.1 depicts this situation. The case without an asymmetry prevention is shown in Figure 4.1a. Figure 4.1b shows the route request and route reply phases if an asymmetry prevention is used. The upper figures show the dissemination of route request messages. The route reply phase is shown in the lower figures.



(a) Without asymmetry prevention **(b)** With asymmetry prevention

Figure 4.1: Schematic representation of the location-based intrusion response with and without asymmetry prevention

4.2 Evaluation

The goal of the evaluation is to compare the location-based intrusion response with adaptive transmission power and with and without asymmetry prevention to the naïve approach presented in Chapter 3. The evaluation is based on a series of simulation studies performed with the (further extended) JiST/MobNet simulation tool [56]. Unless specified otherwise, we use the metrics and configurations described in Section 3.2.

4.2.1 Experimental Design

After preliminary simulations, we defined the parameters and factors used for the evaluation such that the system operates within normal bounds.

Parameters

The parameters used for the evaluation are shown in Table 4.1. The parameters differ from those specified in Table 3.1 for the evaluation of the naïve approach in that we fixed the number of misbehaving nodes to 10, thus considering the worst-case scenario. Only to obtain the baseline, for a defenseless network, we considered a variable number of misbehaving nodes while keeping the total number of nodes fixed at 1000. Further, although we considered a variable time t_{reset} after which quarantined areas are revoked in our simulation studies, we only show the results obtained for $t_{reset=300s}$ in the presentation of the results. $t_{reset} = 300s$ showed to be the optimum configuration in all simulation studies we performed.

Table 4.1: Network parameters as used in the simulation studies

Number of nodes	1000
Benign nodes	990
Black hole nodes	10
Quarantine time t_{reset}	$300s$

Factors

The factors that showed to have a considerable influence on our evaluation metrics are (1) the size of quarantined areas as defined by r_{quar} and (2) the velocity of nodes. The values used for evaluation are listed in Table 4.2. For the evaluation of the location-based intrusion response with and without adaptive transmission power and asymmetry prevention, we performed a full-factorial simulation study regarding quarantine radius and velocity of nodes. For the baseline scenarios in a defenseless network, we performed a full-factorial study regarding number of misbehaving nodes and velocity of nodes.

Table 4.2: Factors as used in the simulation studies

Number of misbehaving nodes out of all nodes (considered only for a defenseless network, otherwise fixed to 10)	1, 3, 5, 10
Quarantine radius r_{quar} in meters (note that for the naïve approach r_{quar} is added to the minimal required quarantine radius of 250 meters)	0, 5, 15, 50
Velocity of nodes in meters per second	0, 1, 10

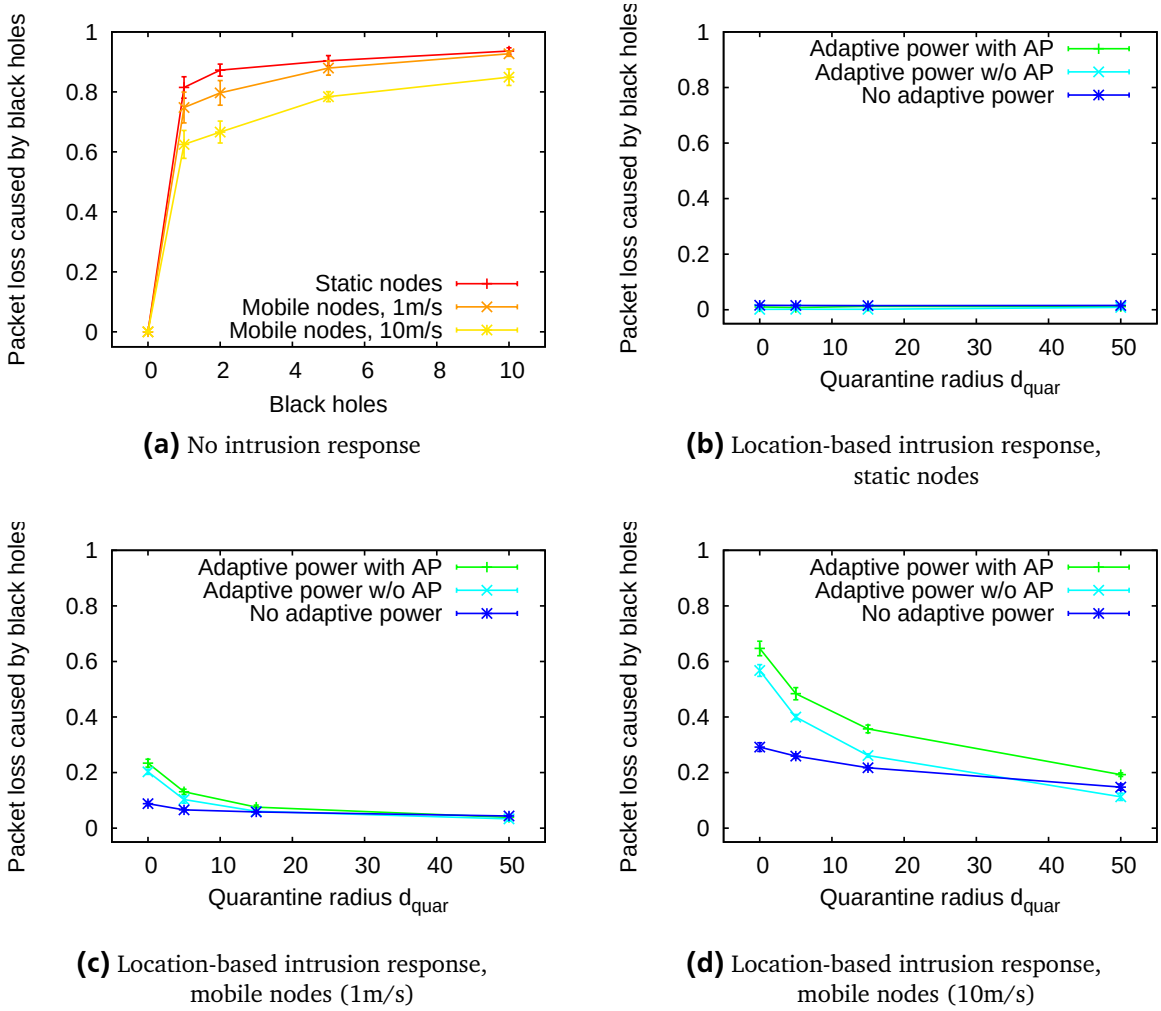
4.2.2 Analysis of the Results

We now discuss the results of the simulation study, proceeding as for the evaluation of the naïve approach for location-based intrusion response presented in Section 3.2.9. In the discussion, we focus on effects caused by an adapted transmission power and an adapted quarantine radius as well as on effects of node velocity. All of the following plots are shown with 95% confidence intervals. By 'AP', we denote the asymmetry prevention.

Packet Loss Caused by Black Holes

The loss caused by black holes in a defenseless network subject to the number of black holes and to the velocity of nodes is shown in Figure 4.2a. We observe that black holes have a higher impact in

scenarios with low node velocity. This can be explained with the help of the length of the routes that are established successfully as shown in Figure 4.12a. Also for scenarios without black hole nodes, the route length decreases if node velocity increases. Thus, the higher the velocity, the shorter the functional routes, the lower the probability for a black hole being part of a route.



(a) No intrusion response

(b) Location-based intrusion response, static nodes

(c) Location-based intrusion response, mobile nodes (1m/s)

(d) Location-based intrusion response, mobile nodes (10m/s)

Figure 4.2: Packet loss caused by black holes

Figures 4.2b to 4.2d show the ratio of packets dropped by black hole nodes subject to the different strategies of adaptive transmission power and subject to the radius of quarantined areas r_{quar}. For static scenarios, we manage to reduce the loss caused by black holes from more than 90% in a defenseless network down to less than 2%. Since we assume that a black hole can not be tracked while it is quarantined, an adaptation of quarantined areas if the black hole moves is not possible. Thus, if nodes move, black holes leave quarantined areas and can become active again. The greater the velocity, the faster this happens. Therefore, we obtain higher loss rates caused by black holes if we increase node velocity. For a velocity of 10 meters per second, also the size of quarantined areas has a remarkable effect. The larger the quarantined area, the longer the protection is effective (recall that for the naïve approach, without adaptive transmission power, the transmission range of 250 meters is added to r_{quar}). If the quarantine radius is chosen too small, a black hole manages to leave quarantine already during the route discovery process and can become active again faster compared to using a larger quarantine radius.

What stands out in the scenario with a node velocity of 10 meters per second is that the loss caused by black holes is slightly but statistically significantly higher for an adaptive transmission power with asymmetry prevention. Again, this can be explained by the route length observed in this scenario as

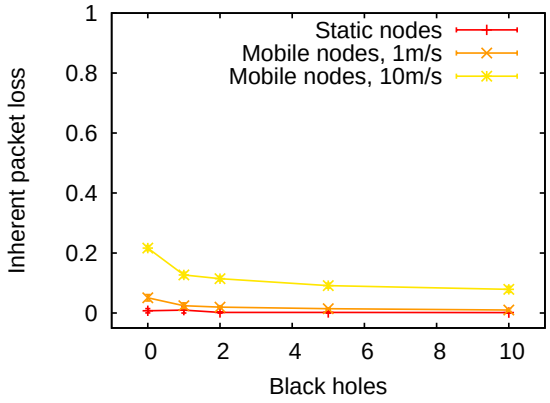
shown in Figure 4.12d. Since the route length is higher if an adaptive transmission power with asymmetry prevention is used, the probability for a black hole being part of a route is increased compared to the other strategies.

Altogether, by using adaptive transmission power, we manage to considerably reduce the size of quarantined areas. Still, we are able to keep the packet loss caused by black holes comparable to the naïve approach of location-based intrusion response for appropriate choices of the size of quarantined areas with respect to node velocity.

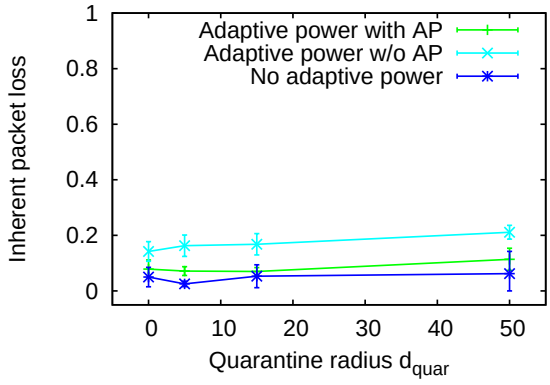
Packet Loss Caused by Inherent Network Properties

Figure 4.3 shows the packet loss caused by inherent network properties which is the sum of (1) the loss caused by collisions on the wireless medium shown in Figure 4.4, (2) the loss caused by exhausted queuing space in the network interface shown in Figure 4.5, and (3) the loss caused by the AODV routing protocol in case of route breaks shown in Figure 4.6.

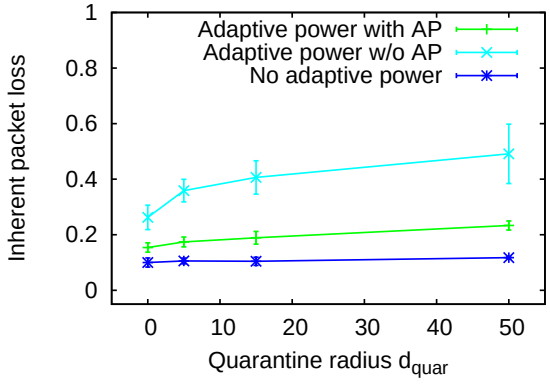
We observe that the inherent loss is strongly affected by both node velocity and the different strategies for location-based intrusion response. For all strategies, the inherent loss increases along with node velocity. For all node velocities, the loss is higher if an adaptive transmission power is used; highest for an adaptive transmission power without asymmetry prevention. The largest part of this loss is caused by the AODV routing protocol, as we will see in the following.



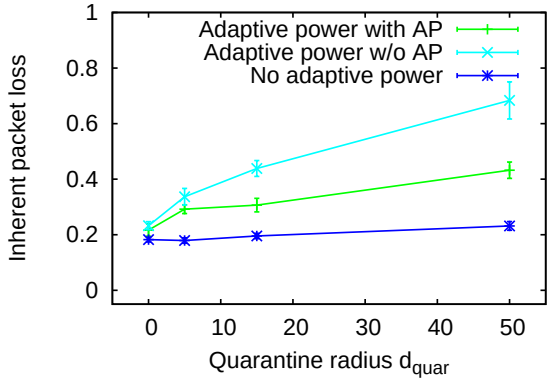
(a) No intrusion response



(b) Location-based intrusion response, static nodes



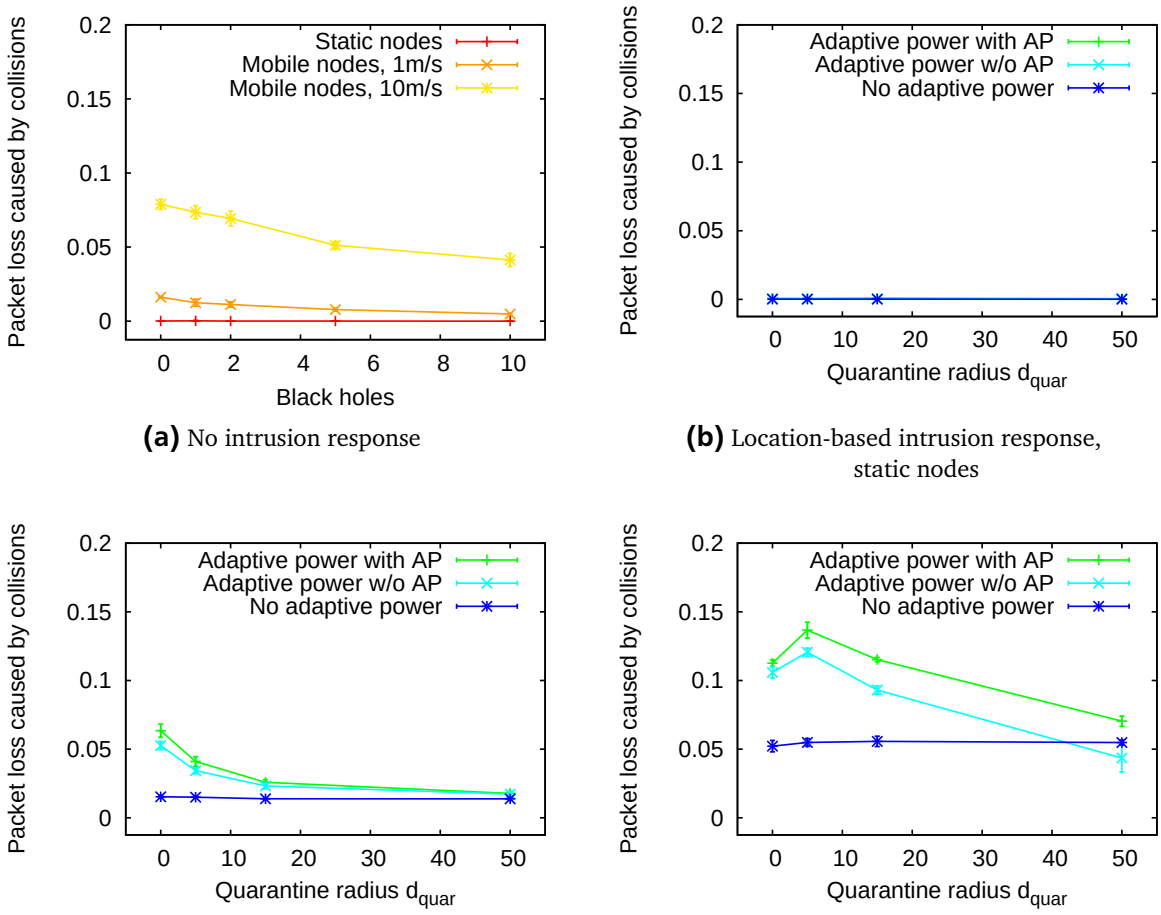
(c) Location-based intrusion response, mobile nodes (1m/s)



(d) Location-based intrusion response, mobile nodes (10m/s)

Figure 4.3: Packet loss caused by inherent network properties

The loss caused by collisions on the wireless medium is shown in Figure 4.4. For a defenseless network, the effects are comparable to those observed in Section 3.2.9 and can be explained in analogy. For scenarios with location-based intrusion response and mobile nodes, we observe an increased loss for small-sized quarantined areas if an adaptive transmission power is used. This can be explained together with the increased black hole activity for small-sized quarantined areas observed in Figure 4.2 by reasons similar to those given in Section 3.2.9. Further, the number of packets dropped due to collisions increases along with the velocity of nodes. This can be explained by the packet loss caused by AODV due to route breaks, shown in Figure 4.6. The probability of route breaks increases along with the velocity of nodes. An increased number of route breaks leads to an increased number of route requests. Since these are sent as broadcast messages, we obtain an increased network load which leads to an increased number of collisions.



(a) No intrusion response

(b) Location-based intrusion response, static nodes

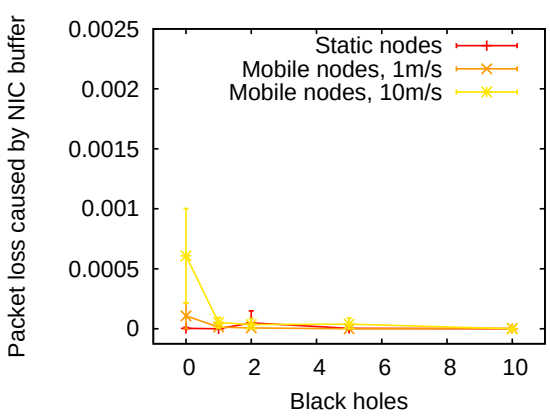
(c) Location-based intrusion response, mobile nodes (1m/s)

(d) Location-based intrusion response, mobile nodes (10m/s)

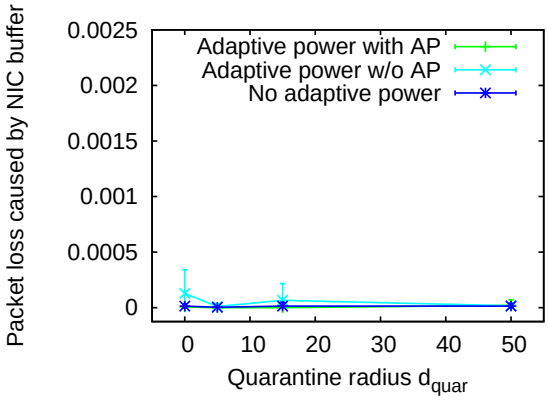
Figure 4.4: Packet loss caused by collisions on the wireless medium

The packet loss caused by exhausted queuing space shows only minor variations as depicted in Figure 4.5. The tendency towards a higher loss for larger quarantined areas and adaptive power observed in Figure 4.5d can be explained by the corresponding loss caused by AODV due to route breaks shown in Figure 4.6d. If the number of route breaks increases, the number of packets that have to be queued and, potentially, dropped if queuing space is exhausted until a new route is discovered also increases.

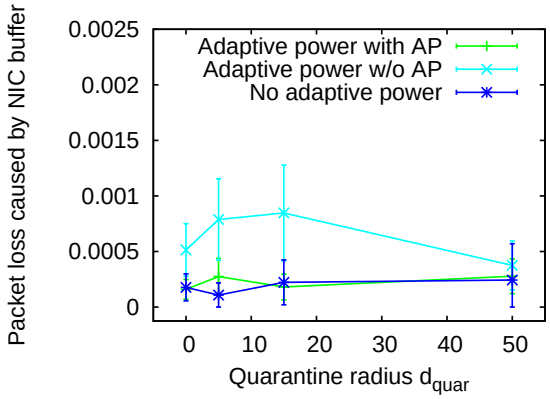
Figure 4.6 shows the packet loss caused by the AODV routing protocol. Here, clearly the drawbacks of an adaptive transmission power become obvious. Without the asymmetry prevention strategy, even for a low node velocity and static scenarios, the loss caused by the routing protocol is intolerable. The effects



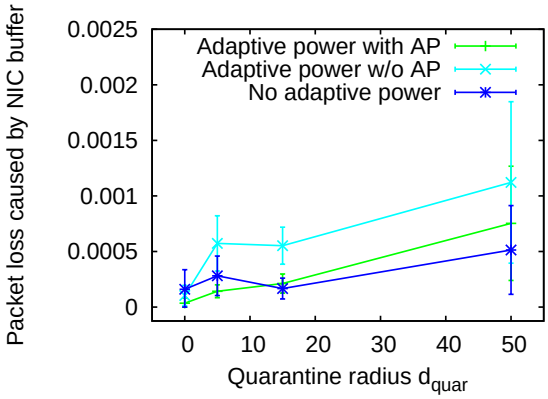
(a) No intrusion response



(b) Location-based intrusion response, static nodes



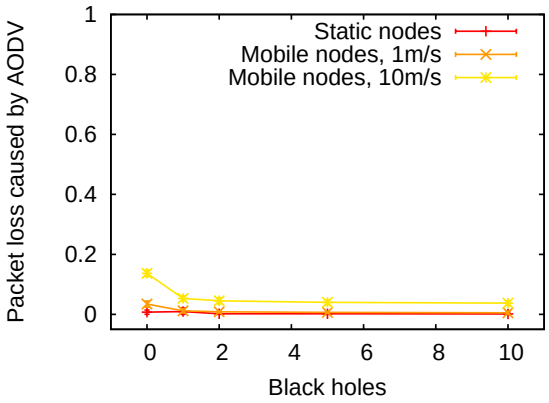
(c) Location-based intrusion response, mobile nodes (1m/s)



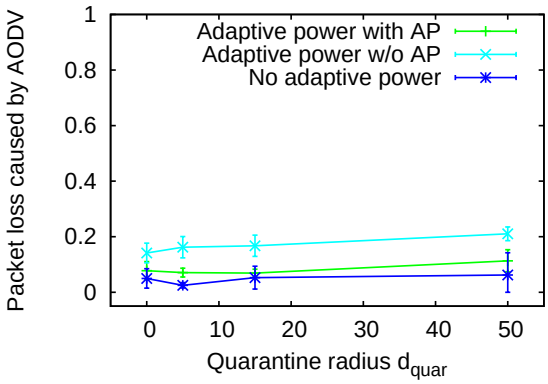
(d) Location-based intrusion response, mobile nodes (10m/s)

Figure 4.5: Packet loss caused by exhausted queuing space in the network interface

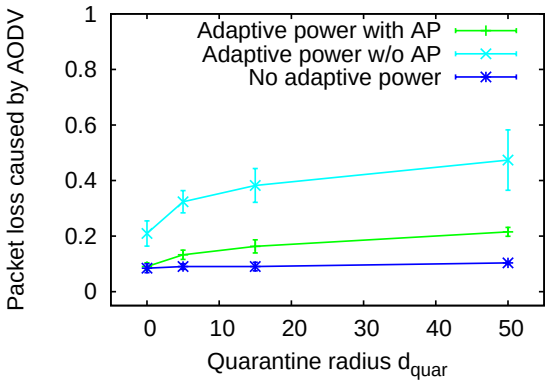
that we depicted schematically in Figure 4.1 lead to a packet loss reaching from 20% for static scenarios up to more than 60% in scenarios with a node velocity of 10 meters per second. Considerably better, but still worse results than for the naïve approach approach without adaptive power can be achieved with the asymmetry prevention strategy. We can conclude that a reduced transmission power leads to an increased loss due to broken routes or routes that cannot be established at all, since we have less room for node movement without leaving transmission ranges. The effect is amplified if the quarantine radius is increased, since the transmission power in this case has to be reduced even further.



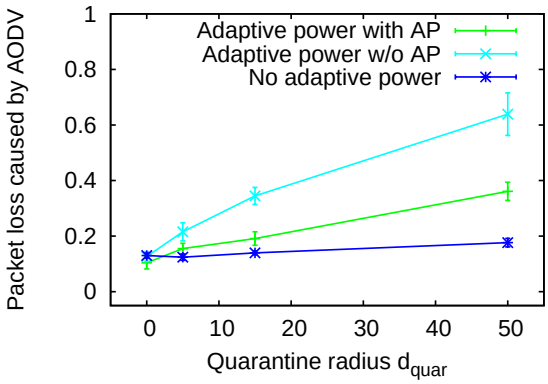
(a) No intrusion response



(b) Location-based intrusion response, static nodes



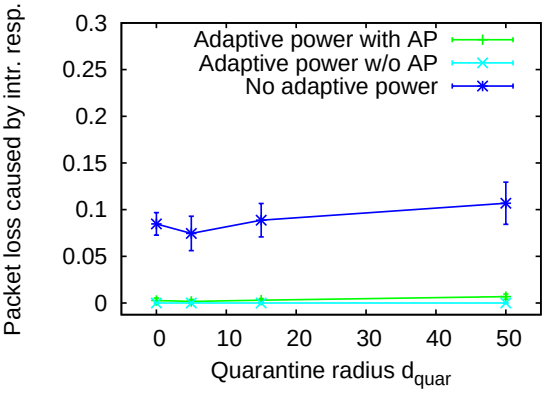
(c) Location-based intrusion response, mobile nodes (1m/s)



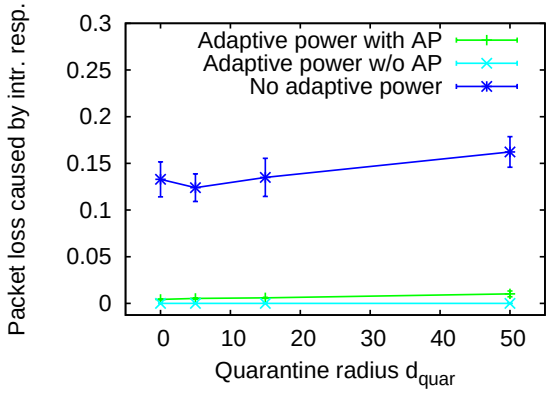
(d) Location-based intrusion response, mobile nodes (10m/s)

Figure 4.6: Packet loss caused by the AODV routing protocol

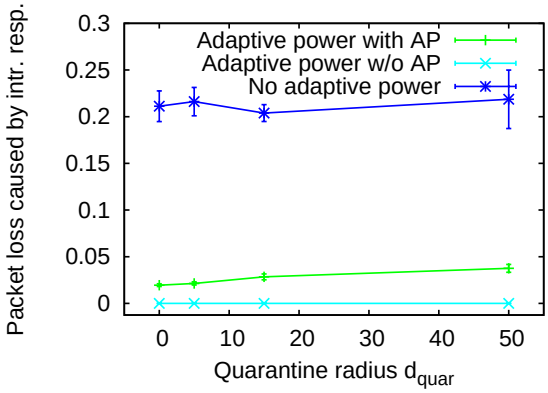
The packet loss caused by intrusion response is shown in Figure 4.7. The results show that, with the power adaptive strategies, we achieve our goal of reducing the loss caused by intrusion response itself. Even for a high node velocity, we achieve a loss of less than 5% which can be explained by the reduced size of quarantined areas and, along with this, by the reduced number of benign nodes affected. It stands out that the loss caused by intrusion response is always (nearly) zero for an adaptive power without asymmetry prevention. This is due to the fact that, in this case, the packet loss is shifted to the AODV routing protocol for which we observed loss rates above 60% because routes break frequently or cannot be discovered at all.



(a) Location-based intrusion response, static nodes



(b) Location-based intrusion response, mobile nodes (1m/s)



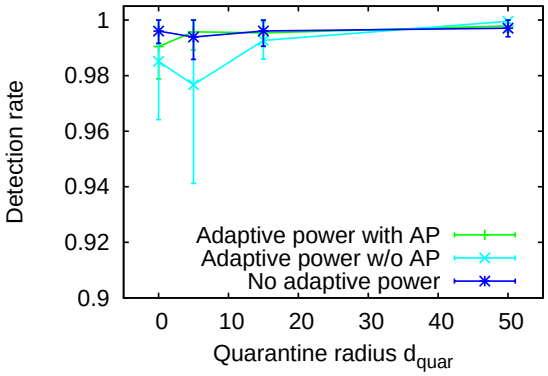
(c) Location-based intrusion response, mobile nodes (10m/s)

Figure 4.7: Packet loss caused by intrusion response

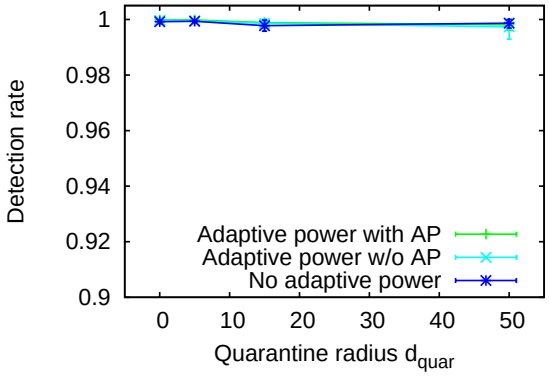
To ensure that equal preconditions regarding the intrusion detection system hold for all variants of location-based intrusion response and adaptive transmission power, we show the detection rate and the false positive rate of the intrusion detection system in Figures 4.8 and 4.9. The results do not show unexpected anomalies except for the large confidence intervals for an adaptive power without asymmetry prevention observed for scenarios with static nodes and small-sized quarantined areas for both detection performance and false positives.

The variations in detection performance directly depend on the increased false positive rate and can be explained in analogy to the reasons provided in Section 3.2.9. In case of a large number of false positive detections, the black hole may be already isolated without being detected directly. This rare case decreases the detection rate slightly.

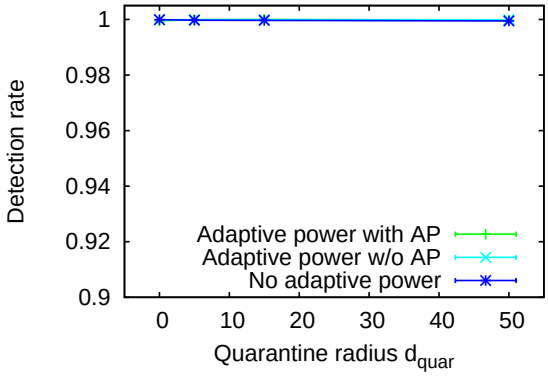
The increased false positive rate and its large confidence intervals are a direct result of the asymmetry effect described in Section 4.1.3. Since the intrusion detection system monitors all unicast messages (that is, it does not differentiate between route reply messages and data messages), route reply messages that are received, but not forwarded correctly may lead to a false positive detection of a benign node. The effect is increased for small-sized quarantine areas, since, in this case, the transmission power may be reduced strongly, for nodes being located in close proximity to misbehaving nodes. Thus, the further transmission power is reduced, the more likely asymmetry effects will occur.



(a) Location-based intrusion response, static nodes

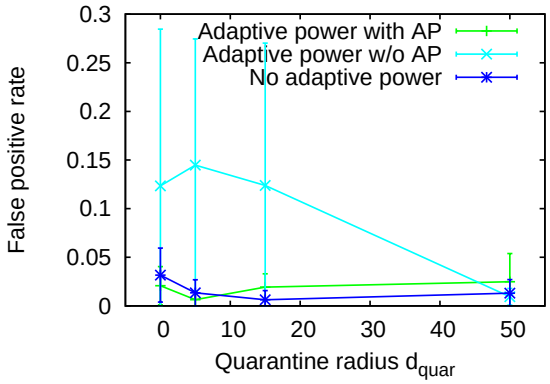


(b) Location-based intrusion response, mobile nodes (1m/s)

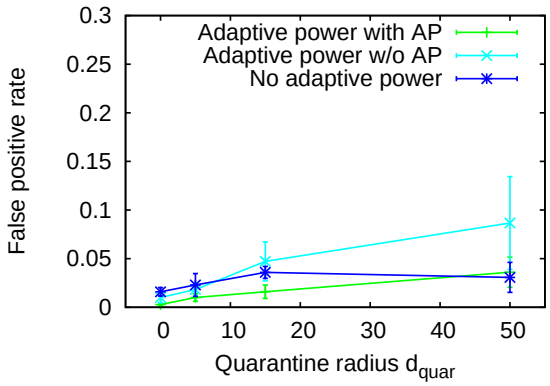


(c) Location-based intrusion response, mobile nodes (10m/s)

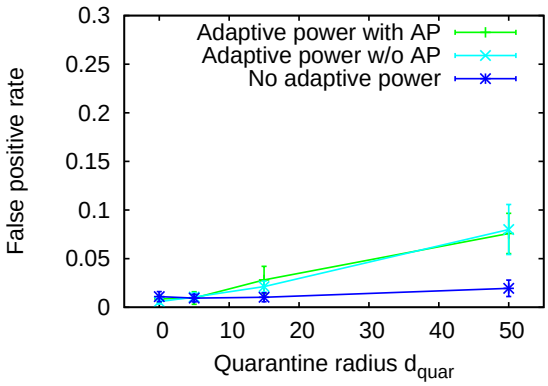
Figure 4.8: Detection rate of the intrusion detection system



(a) Location-based intrusion response, static nodes



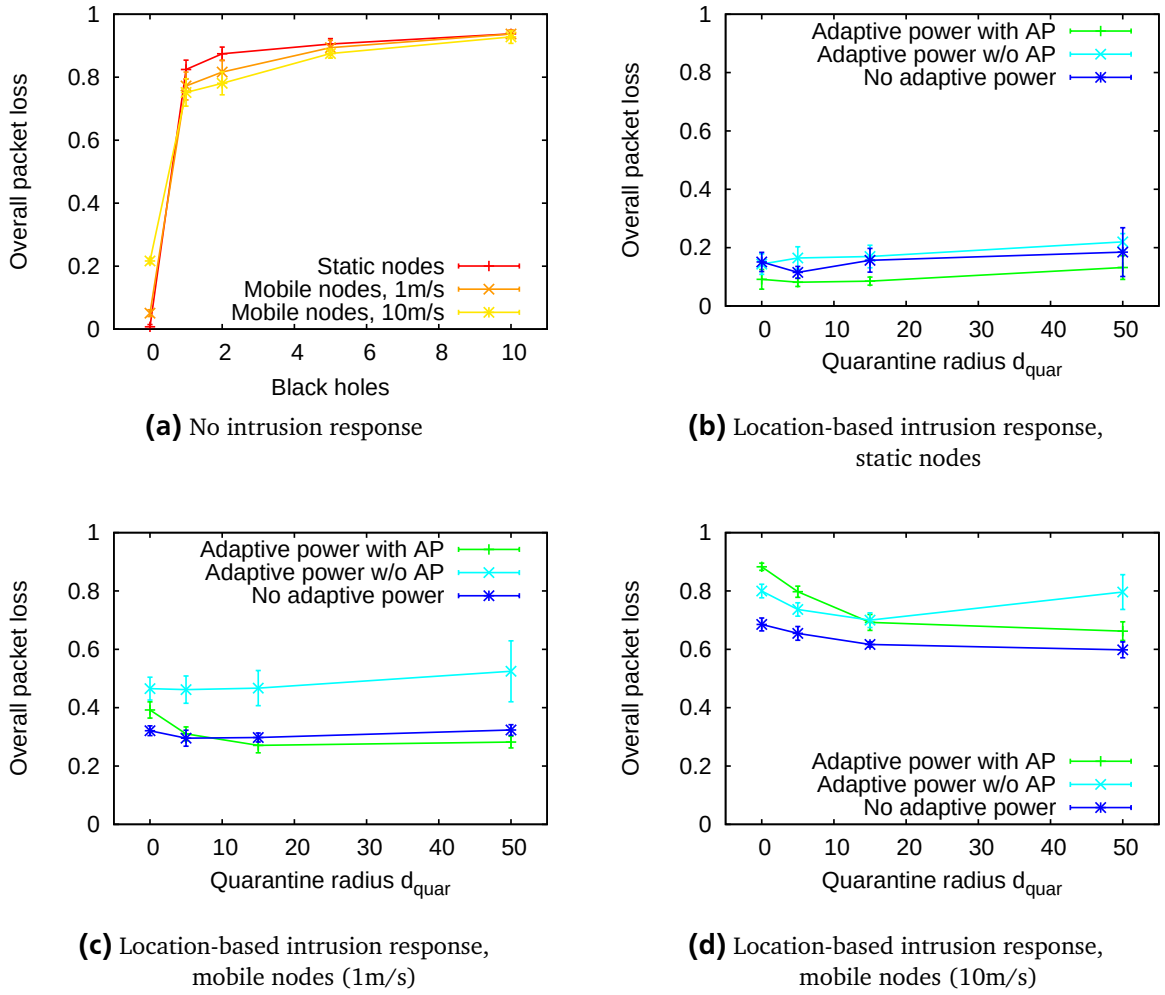
(b) Location-based intrusion response, mobile nodes (1m/s)



(c) Location-based intrusion response, mobile nodes (10m/s)

Figure 4.9: False positive rate of the intrusion detection system

The overall packet loss observed in the network is shown in Figure 4.10. Here, it gets obvious that the improvement of a significantly reduced loss caused by intrusion response is canceled out by the increased loss caused by AODV. A minor overall improvement can be observed for a low velocity, if an adaptive power is used together with asymmetry prevention. Without asymmetry prevention, the adaptive power performs worse than the naïve approach for location-based intrusion response without adaptive power.



(a) No intrusion response

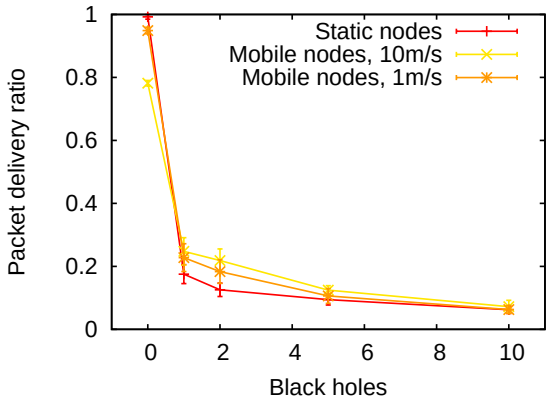
(b) Location-based intrusion response, static nodes

(c) Location-based intrusion response, mobile nodes (1m/s)

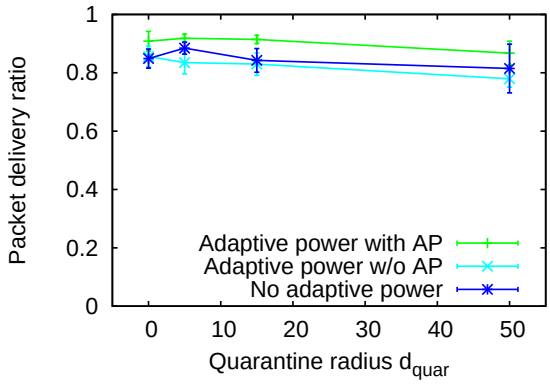
(d) Location-based intrusion response, mobile nodes (10m/s)

Figure 4.10: Overall packet loss

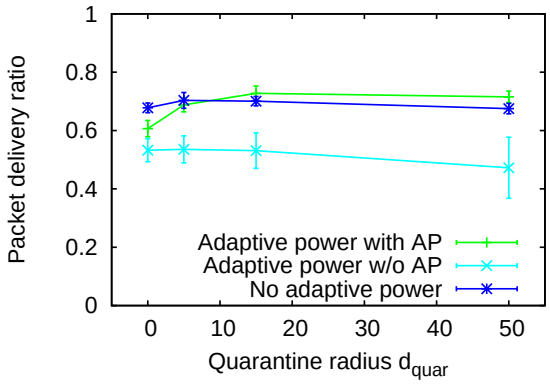
To double check the correctness of the packet loss, we present the overall delivery ratio in Figure 4.11 (recall that packet delivery is monitored independently from packet loss). Since the delivery ratios match the overall packet loss, we assume correctness of the results presented so far.



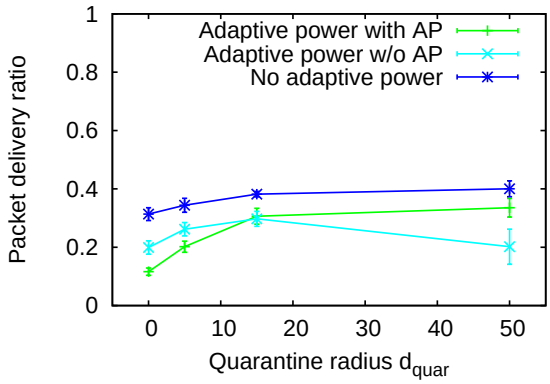
(a) No intrusion response



(b) Location-based intrusion response, static nodes



(c) Location-based intrusion response, mobile nodes (1m/s)



(d) Location-based intrusion response, mobile nodes (10m/s)

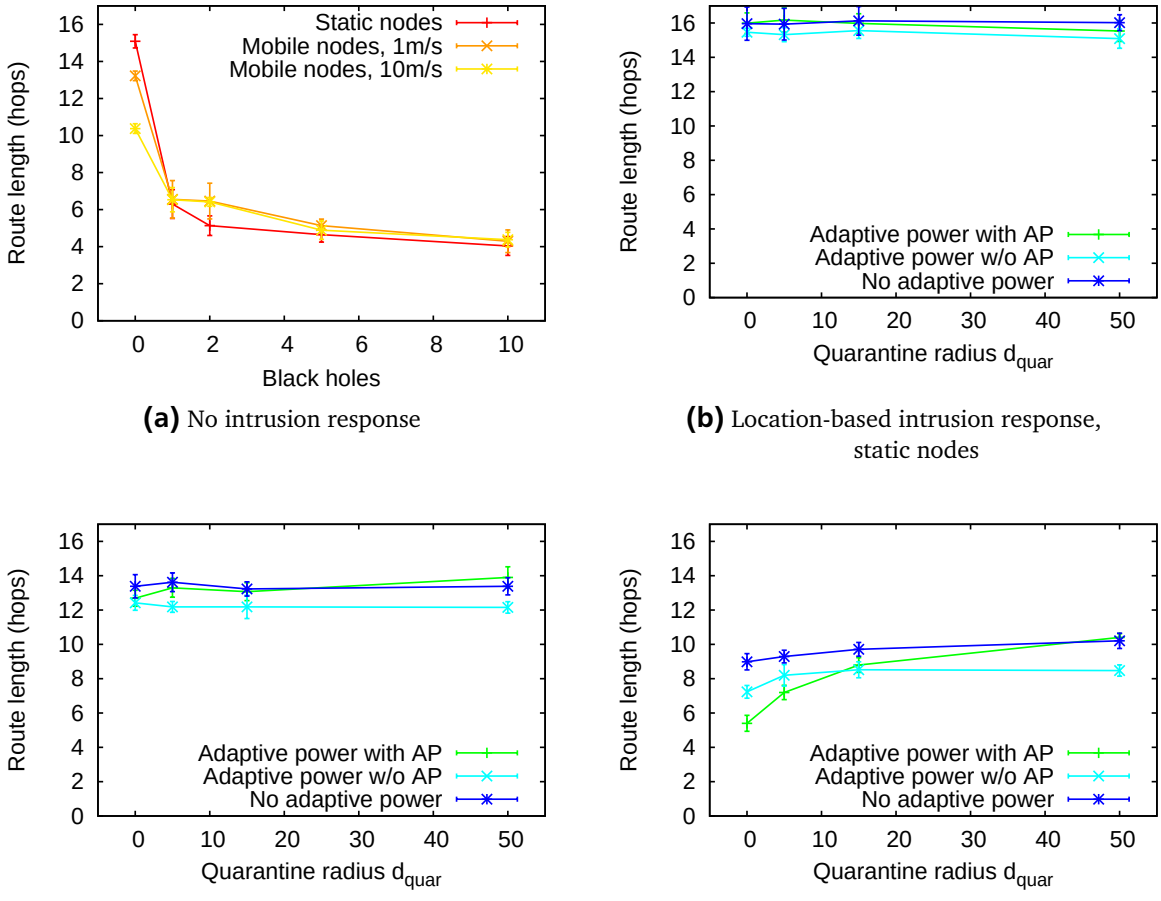
Figure 4.11: Overall delivery ratio

Route Lengths

Figure 4.12 shows how the length of routes that have been established successfully is affected by the velocity of nodes as well as by the different strategies for location-based intrusion response. For all strategies, the route length decreases when node velocity increases. Thus, node mobility not only causes route breaks but also affects the route discovery process, which only takes a relatively short time. Obviously, during the time a route request is propagated to the destination until the route reply is sent back, the topology changes so rapidly that the path via which the route request reached the destination is not available anymore.

Regarding the different strategies for location-based intrusion response, we observe that routes tend to be shorter if an adaptive transmission power is used in scenarios with high node velocity. This can be explained in analogy to the reasons provided in Section 3.2.9. Recalling that route lengths are monitored on a per packet basis, the probability for route breaks increases if transmission power decreases because nodes move out of each others transmission range faster. Thus, less packets are transmitted successfully over long routes if transmission power is reduced.

For scenarios with high node velocity, we also observe a shorter route length for small-sized quarantined areas, if an adaptive transmission power is used. Since black holes moving at a high velocity leave small-sized quarantined areas after a short time, black hole activity is increased as shown in Figure 4.2d. Thus, route length decreases in this case.



(a) No intrusion response

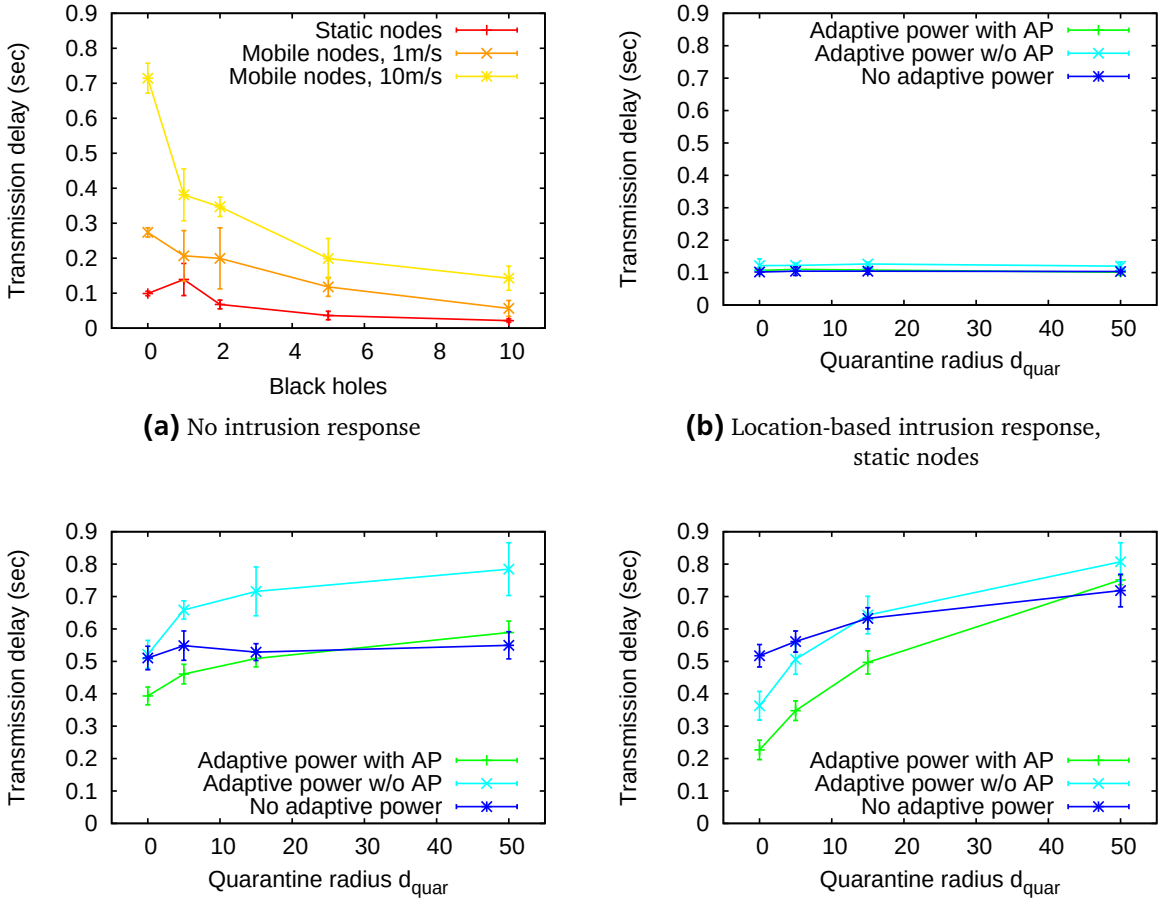
(b) Location-based intrusion response, static nodes

(c) Location-based intrusion response, mobile nodes (1m/s)

(d) Location-based intrusion response, mobile nodes (10m/s)

Figure 4.12: Route length

The transmission delay measured from sending application to receiving application is shown in Figure 4.13. About 90% of the delay observed from application to application is caused by the routing delay on the source node shown in Figure 4.14. The propagation delay of the multi-hop route shown in Figure 4.15 makes up the remainder.



(a) No intrusion response

(b) Location-based intrusion response, static nodes

(c) Location-based intrusion response, mobile nodes (1m/s)

(d) Location-based intrusion response, mobile nodes (10m/s)

Figure 4.13: Transmission delay from application to application

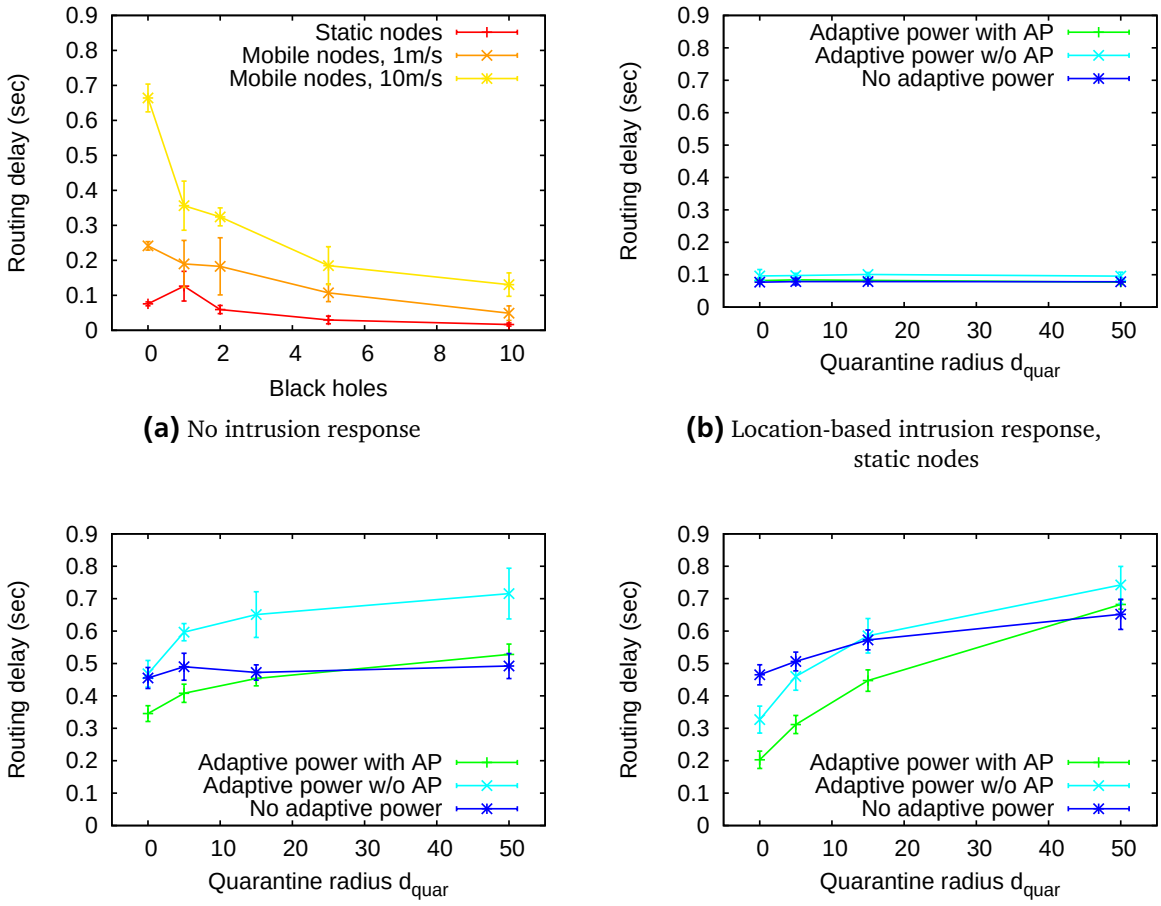
We observe that the transmission delay is strongly affected by node velocity. This is explained by the fact that routes are stable only for a short time if nodes move at high velocity. Thus, if node velocity increases, new routes have to be discovered frequently, causing an increased routing delay as a major part of the transmission delay.

Also for the propagation delay it holds that delay increases along with node velocity. This is due to the fact that network traffic increases along with node velocity due to the increased route maintenance required. If network traffic increases, more packets have to be queued before transmission, which also can be deduced from the increased loss due to exhausted queuing space shown in Figure 4.5. Thus, also the propagation delay of a multi-hop route increases.

Regarding the variants of location-based intrusion response with adaptive power, we observe a shorter delay for small-sized quarantined areas which can be explained in analogy to the reasons given in Section 3.2.9. Small sized-quarantined areas lead to an increased activity of black hole nodes and, thus, to an increased number of route breaks induced by intrusion response. Since, in case of a route break, the

expanding ring search of AODV continues where it was left off and not from the beginning, the average time required for route discovery decreases in this case.

Further, the delay is higher if an adaptive power without asymmetry prevention is used in scenarios with mobile nodes, which is best seen in Figure 4.13c, where the effect is only marginally mixed with effects of black hole activity. This is due to asymmetry effects causing a higher number of route discovery processes required to establish a functional route, amplified by route breaks induced by node mobility.



(a) No intrusion response

(b) Location-based intrusion response, static nodes

(c) Location-based intrusion response, mobile nodes (1m/s)

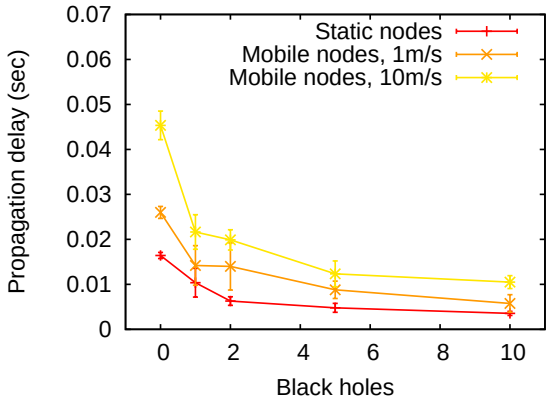
(d) Location-based intrusion response, mobile nodes (10m/s)

Figure 4.14: Routing delay on source node

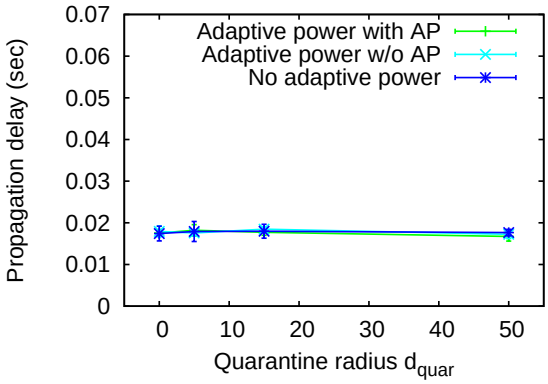
4.3 Conclusion - Adaptive Transmission Power

In this chapter, we identified possibilities and limitations of supporting location-based intrusion response in mobile ad hoc networks by an adaptive transmission power. We proposed a simple power reduction scheme and a scheme that prevents effects of asymmetric links that arise when using an adaptive power in combination with the IEEE 802.11 medium access control scheme.

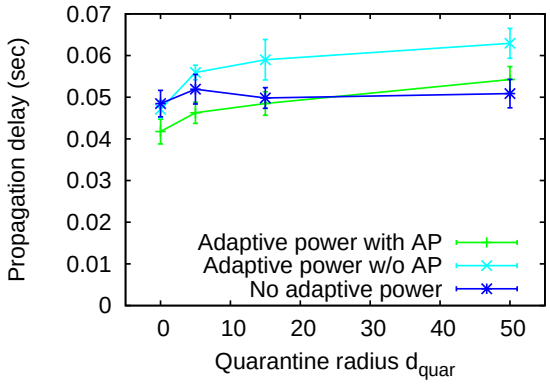
The evaluation showed that an adaptive power is able to considerably improve unwanted side-effects of the location-based intrusion response. Both the size of quarantined areas and the packet loss caused by intrusion response itself were reduced significantly. Still, the approach suffers from an increased packet loss caused by the (non power-aware) AODV routing protocol. We conclude that an adaptive transmission power can be used to effectively support location-based intrusion response in mobile ad hoc networks, if the routing protocol deployed is able to handle the resulting effects.



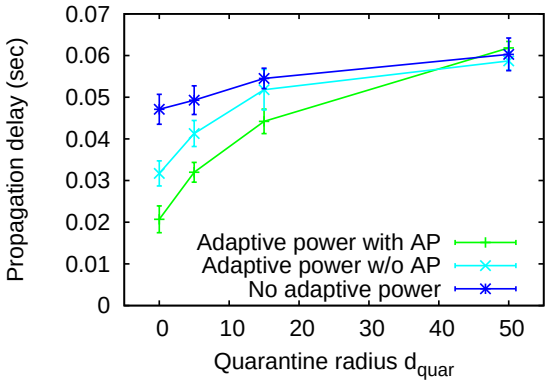
(a) No intrusion response



(b) Location-based intrusion response, static nodes



(c) Location-based intrusion response, mobile nodes (1m/s)



(d) Location-based intrusion response, mobile nodes (10m/s)

Figure 4.15: Propagation delay from source to destination

5 Supporting Location-based Intrusion Response in Mobile Ad Hoc Networks with Delay Tolerant Communication

In this chapter, we describe how time can be used as a further dimension for overcoming the limitations of the location-based intrusion response in mobile ad hoc networks stated in Section 3.4. By harnessing delay tolerance of applications such as e-mail, we increase network performance in terms of packets transmitted successfully. We describe different strategies of how a delayed communication can be realized. In a series of simulation studies we scrutinize the trade-offs between reliability and transmission delay.

5.1 Architecture

We now describe how the system architecture for location-based intrusion response presented in Section 3.1 is extended to support delay tolerant communication.

5.1.1 The Bundle Layer

To enable delay tolerant communication, we use a simplified version of the Bundle Layer discussed in Section 2.9. The combination with the Licklider Protocol is not necessary since, in our scenario, the wireless data link layer is based on the IEEE 802.11 protocol and, thus, is not subject to high delays. Regarding disruption tolerance, also a proxied TCP variant would have been applicable. Yet, for reasons of flexibility, we chose not to be bound to TCP as transport protocol.

The mode of operation of our implementation of the Bundle Protocol is shown in Figure 5.1. The Bundle Layer is added between transport layer and application layer as specified in [94]. Packets that would have been dropped by the location-based intrusion response due to quarantined benign nodes are handed to the Bundle Layer for delayed retransmission. This is done at the sender as well as at each intermediate node.

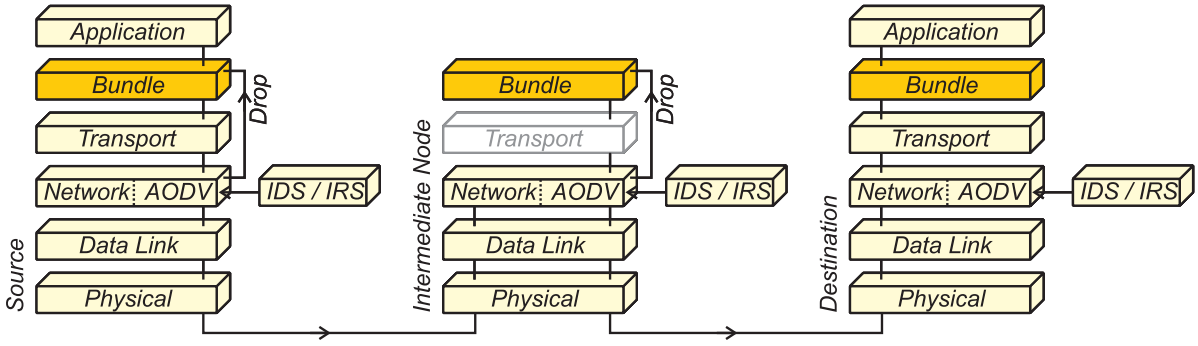


Figure 5.1: Integration of the Bundle Layer for managing delay tolerant communication

The Bundle Layer is parameterized by the time span $t_{retrans}$ between two subsequent retransmission attempts and the number $n_{retrans}$ of retransmission attempts that are performed per packet. Note that $n_{retrans}$ is a global value. The number of retransmission attempts left is (besides other information required) transmitted in an additional Bundle Header of the packet.

5.1.2 Buffered and Unbuffered Intrusion Detection

Our intrusion detection system, as described in Section 3.1.4, is based on monitoring the packet forwarding behavior of nodes in transmission range. If a monitored node turns out to be a black hole, all packets sent to the black hole during the monitoring phase are lost.

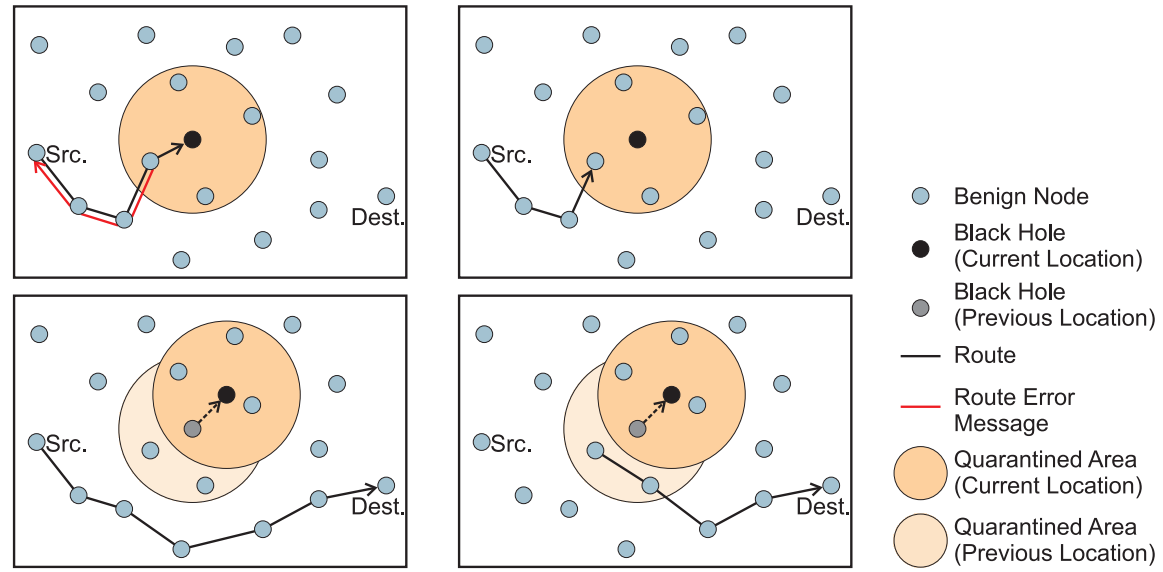
To mitigate this drawback, we extend the intrusion detection system by a buffered monitoring phase. In contrast to the first, unbuffered version described in Section 3.1.4, all packets that are sent during the monitoring phase are buffered by the monitoring node. If the monitored node is classified as a black hole, the buffer content is handed to the Bundle Layer for retransmission. Otherwise, the monitoring node flushes the buffer without further action.

5.1.3 Transparent and Non-transparent Location-based Intrusion Response

As described in Section 3.1.6, the location-based intrusion response mechanism excludes a misbehaving node based on its location. If a node B is classified as a black hole by a node X, X establishes a quarantined area at B's location.

In the first version of the location-based intrusion response, if X is located in a quarantined area, it interrupts all active route which it is part of. If X is part of a route from a source S to a destination D, X informs S by sending a corresponding route error message. Thus, intrusion response is non-transparent in that S is aware of the route break and is in duty to find a new route to D, as shown in Figure 5.2a. Only packets that are in flight or still sent by S until the route error message arrives are handed to the Bundle Layer of X for later retransmissions.

For the delay tolerant extension of the location-based intrusion response mechanism, we, additionally, introduce a transparent alternative in which X does not inform S that the route is affected by a misbehaving node. Thus, the route is not interrupted and X hands all packets arriving from S to the Bundle Layer for retransmission. In the transparent case, X is fully responsible for finding a new route to D and retransmitting packets appropriately as soon as it has left the quarantined area, as shown in Figure 5.2b.



(a) Non-transparent intrusion response **(b)** Transparent intrusion response

Figure 5.2: Schematic representation of transparent and non-transparent intrusion response

5.2 Evaluation

The main goal of the evaluation is to quantify the trade-off between transmission delay and packet delivery ratio. We show a comparison of the performance of the location-based intrusion response with and without being supported by the Bundle Layer. We consider the four possible combinations of transparent/non-transparent intrusion response and buffered/unbuffered intrusion detection. The evaluation is based on a series of simulation studies performed with the (further extended) JiST/MobNet simulation tool [56]. Unless specified otherwise, we use the metrics and configurations described in Section 3.2.

5.2.1 Metrics for the Evaluation

To identify the effects caused by harnessing delay tolerance, we modify the metrics of overall packet loss, overall delivery ratios, and overall transmission delay with respect to the definition in Section 3.2.3. For the evaluation of the delay-tolerant approach, we independently monitor the results with and without retransmissions of the Bundle Layer.

5.2.2 Experimental Design

Based on preliminary simulations, we defined the parameters and factors used for the evaluation such that the system operates within normal bounds.

Parameters

The parameters used for the simulation studies that differ from those used for the evaluation of the first approach for location-based intrusion response in Section 3.2.8, are defined in Table 5.1. As for the evaluation of the adaptive transmission power in the previous chapter, we fixed the number of misbehaving nodes to 10, thus considering the worst-case scenario of the evaluation of the first version of location-based intrusion response. Since we observed congestion situations in some scenarios, which we discuss in detail in the following, we reduced the network load to 10 streams running simultaneously. Again, the quarantine time is fixed to $t_{reset} = 300s$ which turned out to be the optimal choice.

Table 5.1: Network parameters as used in the simulation studies

Total number of nodes	1000
Benign nodes	990
Black hole nodes	10
Traffic pattern	Constant bitrate traffic with streams of 2048 bytes per second split in 4 packets per second. 10 streams run in parallel with a duration of 30 seconds each. Streams are set up uniformly within the first 30 seconds of simulated time. After this, every time a stream ends, a new one is created immediately. Source and destination are randomly selected from the set of benign nodes, resulting in that malicious nodes are not chosen as source or destination.
Quarantine time t_{reset}	$300s$

Factors

The factors that considerably affect system performance are the number $n_{retrans}$ of retransmission attempts performed by the Bundle Layer per packet and the interval $t_{retrans}$ between two consecutive retransmission attempts. The values used in the evaluation are defined in Table 5.2. We performed a full-factorial simulation study for all combinations possible of buffered/unbuffered intrusion detection system and transparent/non-transparent intrusion response.

Table 5.2: Factors as used in the simulation studies

Number $n_{retrans}$ of retransmission attempts by the Bundle Layer	4, 16, 32
Interval $t_{retrans}$ between two consecutive retransmissions	10, 20, 30, 40, 50, 60

5.2.3 Analysis of the Results

We now discuss the results of the simulation studies. All plots are given with 95% confidence intervals. Note that, for reasons of readability, we condensed the results to bar charts for all metrics that were not statistically significantly affected by the factors defined in Table 5.2.

Packet Loss Caused by Black Holes

Figure 5.3 shows the packet loss caused by black holes. Note that this metric includes all packets dropped by black holes, also those that are recovered in case of a buffered intrusion detection system. We decided not to include these retransmissions in this metric to keep the results comparable to those obtained previously for the first version of location-based intrusion response and for the intrusion response supported by adaptive transmission power.

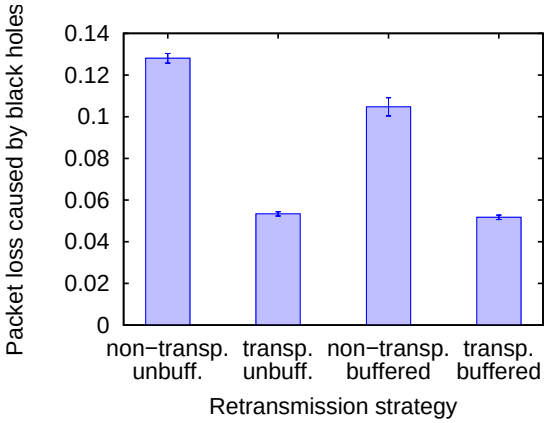


Figure 5.3: Packet loss caused by black holes

The baseline to which we have to compare the results is given by the loss of about 13% observed for an unbuffered intrusion detection system with a non-transparent intrusion response. This combination corresponds to the first version of location-based intrusion response presented in Chapter 3 and the loss observed matches the results obtained previously.

We observe that the loss for the non-transparent intrusion response combined with a buffered intrusion detection is slightly but statistically significantly lower than the baseline. This is due to a slight

overload situation that occurred for this combination, which can be seen from the increased loss caused by exhausted queuing space shown in Figure 5.6; packets that get lost on the way to a black hole due to network congestion cannot be dropped by the black hole itself anymore.

What stands out is that the loss observed for a transparent intrusion response, combined with both buffered and unbuffered intrusion detection, is considerably lower than the baseline and than the loss observed for non-transparent intrusion response with buffered intrusion detection. This can be explained by the increased false positive rate, shown in Figure 5.10, that occurred for these combinations. If false positives occur in proximity to a black hole, the black hole is surrounded by quarantined areas that still prevent the black hole from getting active, although it might have already left the quarantined area that was established at its original location.

Packet Loss Caused by Inherent Network Properties

The packet loss caused by inherent network properties is shown in Figure 5.4 as sum of loss caused by collisions shown in Figure 5.5, loss caused by exhausted queuing space shown in Figure 5.6, and loss caused by the routing protocol due to route breaks or unreachable destinations shown in Figure 5.7. Like the metric of loss caused by black holes, all of these metrics do not consider retransmissions of the Bundle Layer for reasons of comparability. The baseline is again given by the results obtained for unbuffered intrusion detection with non-transparent intrusion response. For this case, the results match those observed in Chapter 3.

Compared to the baseline, the loss caused by inherent network properties is slightly higher for a transparent intrusion response combined with both buffered and unbuffered intrusion detection. This is due to an increased loss caused by the routing protocol for these combinations, as shown in Figure 5.7. This, in turn, can be explained by the increased false positive rate of the intrusion detection system shown in Figure 5.10. An increased false positive rate leads to an increased number of quarantined benign nodes and, thus, to an increased loss due to route breaks or unreachable destinations. As a direct consequence, the loss caused by collisions and the loss caused by exhausted queuing space for a transparent intrusion response is slightly lower than the baseline. If less packets are sent through the network due to unreachable destinations, less collisions occur and queuing load is reduced.

It stands out that the inherent loss for a non-transparent intrusion response with a buffered intrusion detection is considerably higher than the inherent loss observed for all other combinations. From the packet loss caused by exhausted queuing space shown in Figure 5.6, we learn that this is due to a slight overload situation of the network. The overload leads to an increased loss caused by the routing protocol due to route request or route reply messages that are dropped if queuing space is exhausted. As a consequence, routes cannot be established and packet loss occurs due to unreachable destinations.

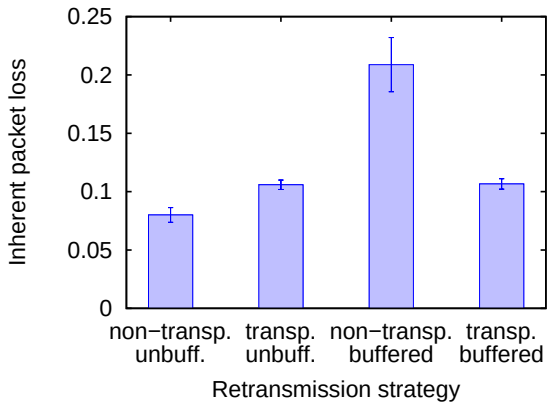


Figure 5.4: Packet loss caused by inherent network properties

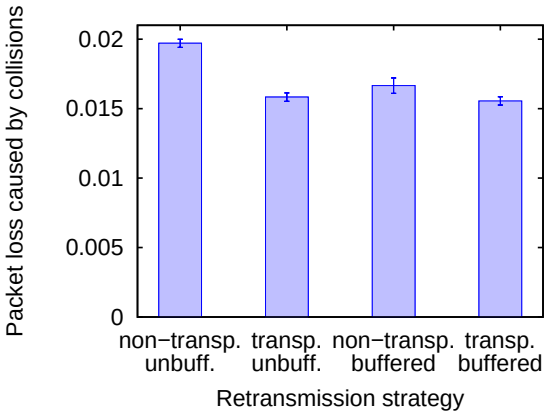


Figure 5.5: Packet loss caused by collisions on the wireless medium

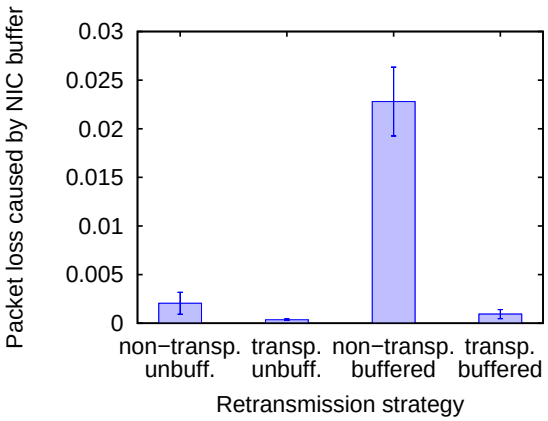


Figure 5.6: Packet loss caused by exhausted queuing space in the network interface

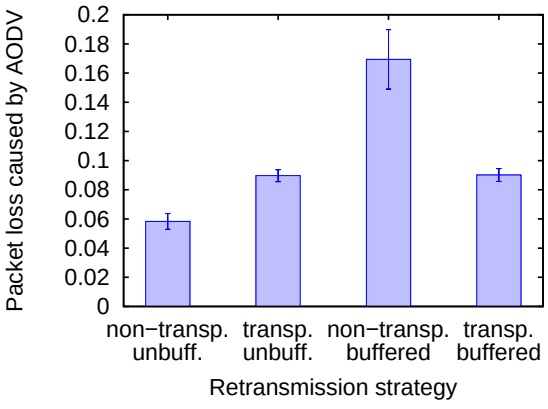


Figure 5.7: Packet loss caused by the AODV routing protocol

The packet loss caused by location-based intrusion response due to quarantined benign nodes is shown in Figure 5.8. The baseline is given again by the loss observed for the combination of non-transparent intrusion response and unbuffered intrusion detection. Retransmissions by the Bundle Layer are not considered.

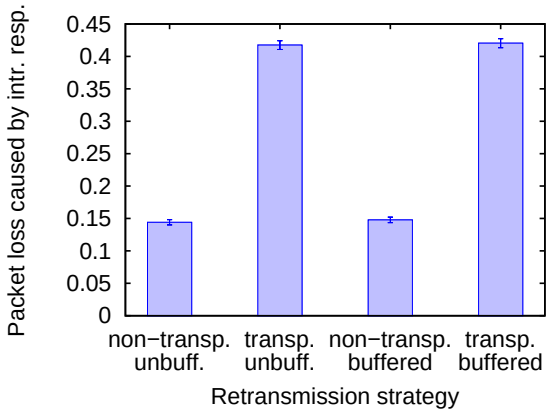


Figure 5.8: Packet loss caused by intrusion response

Regarding detection performance of the intrusion detection system, we did not observe any variations of statistical significance subject to the different combinations of transparent/non-transparent intrusion response and buffered/unbuffered intrusion detection, as shown in Figure 5.9. Therefore, equal preconditions apply for all schemes.

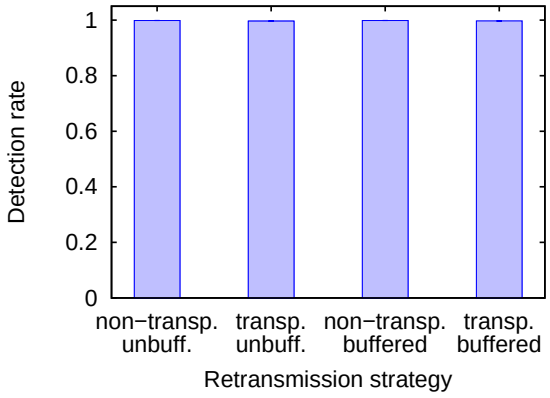
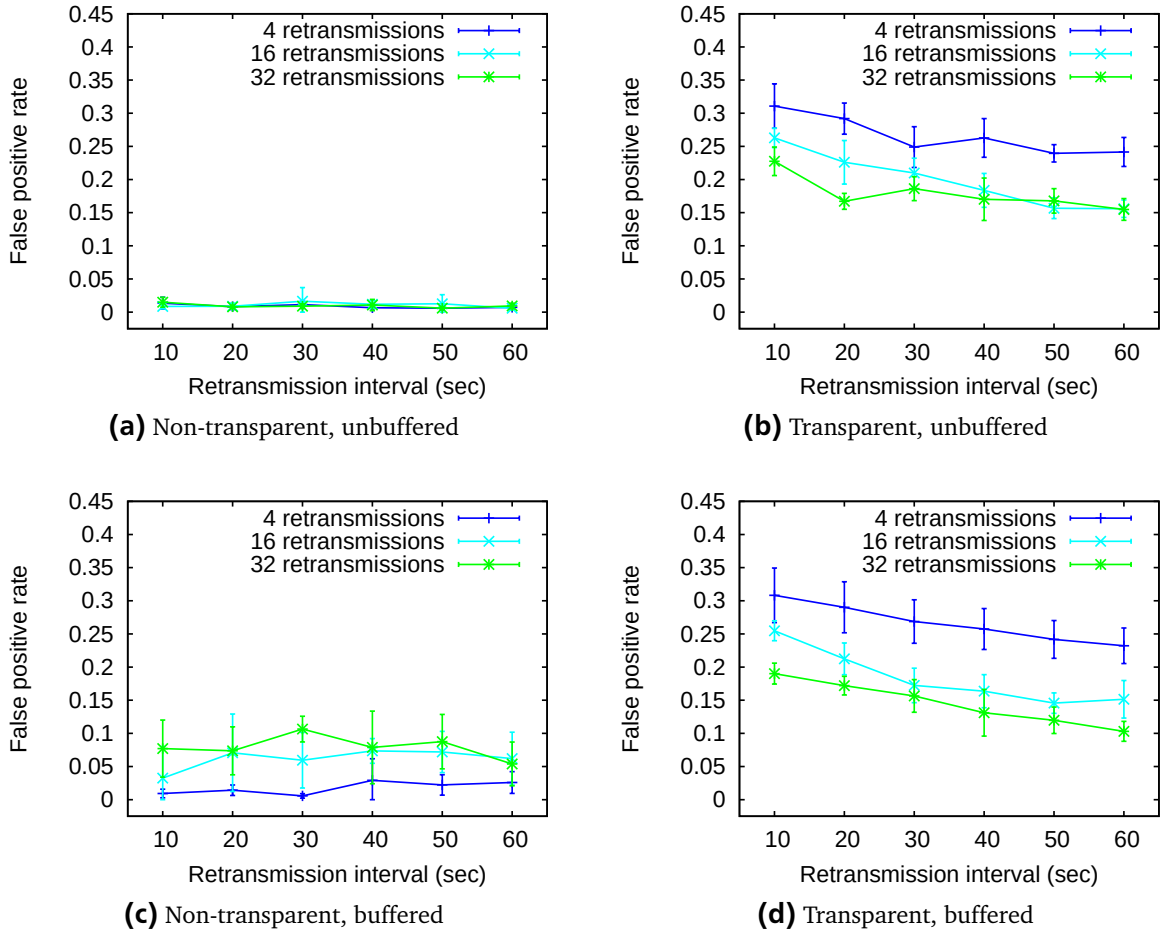


Figure 5.9: Detection rate of the intrusion detection system

We expected a slightly lower loss caused by intrusion response for the non-transparent approach with buffered intrusion detection due to the slight overload caused by this combination. In analogy to the reasons given for the loss caused by black holes, packets that get lost on their way to a quarantined area due to network overload do not affect the loss caused due to the quarantined area itself anymore. Yet, we observe that the loss caused by intrusion response for a non-transparent intrusion response with a buffered intrusion detection does not differ from the baseline. This can be explained by the slightly increased false positive rate for this combination shown in Figure 5.10c. False positives cause an increased number of quarantined areas which, in turn, cause an increased loss due to intrusion response. Altogether, the reduced loss due to network overload and the increased loss due to false positives balance out each other. The slightly increased false positive rate also results from the slight overload situation, which increases packet loss caused by inherent network properties as explained in the previous section.

If this loss exceeds the classification threshold $thres_{black}$ of the intrusion detection system, false positives occur.

Compared to the baseline loss of about 15%, the loss for transparent intrusion response with both buffered and unbuffered intrusion detection is considerably higher at about 40%. The main reason for this is that, in the transparent approaches, the sender is not informed when a route is affected by misbehavior. Instead, all packets are buffered at the first quarantined node. Since we want to maintain comparability of the results, these packets are considered lost before they are retransmitted by the Bundle Layer. Thus, not only packets in flight while a route error message travels back to the sender, but all packets that belong to a stream affected by misbehavior are considered lost in the first step. This effect is further amplified by an increased false positive rate for the transparent intrusion response schemes which can be seen in Figures 5.10b and 5.10d. The increased false positive rate, in case of transparent intrusion response, is a drawback of the simplified intrusion detection system that works based on local information only. No information is exchanged to support monitoring the forwarding behavior of other nodes. Thus, a quarantined benign node that does not forward, but buffers packets, may be classified as black hole by a node that is not aware of the corresponding quarantined area.



(a) Non-transparent, unbuffered

(b) Transparent, unbuffered

(c) Non-transparent, buffered

(d) Transparent, buffered

Figure 5.10: False positive rate of the intrusion detection system

Regarding the false positive rate of the transparent intrusion response schemes shown in Figures 5.10b and 5.10d, it further stands out that the false positive rate decreases if the number of retransmission attempts increases and if the time between two consecutive retransmission attempts increases. We put this down to the fact that, in the transparent schemes, packets are stored and forwarded in the network. If we further consider the routing metric of AODV, which prefers short routes, it follows that packets get closer to their destination with each retransmission attempt. Also, if we increase the time between

consecutive retransmission attempts, the probability for a quarantined benign node having left the quarantined area and, thus, being able to transmit successfully increases. Now, if packets get closer to the destination with increasing number of retransmissions and increasing time between consecutive retransmissions, the probability for a route to the destination being affected by misbehavior decreases due to the expanding ring search of AODV. If the destination is located in a ring closer to the source than the next misbehaving node, the route request will not reach the misbehaving node and the route can be established successfully. Finally, if the probability for routes being affected by misbehavior decreases, the probability for false positive classifications of quarantined benign nodes also decreases.

Overall Packet Loss and Delivery Ratios

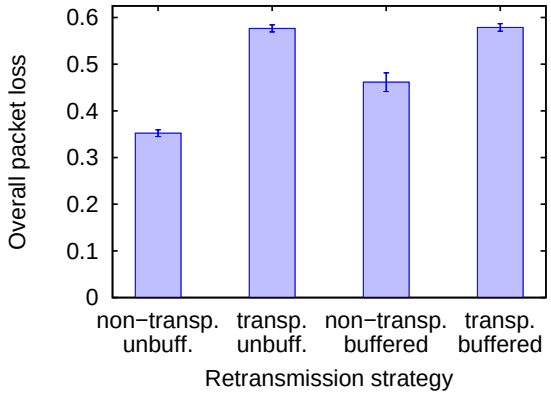


Figure 5.11: Overall packet loss without retransmissions of the bundle layer

The overall packet loss without considering the retransmissions of the Bundle Layer is shown in Figure 5.11. The baseline given by the loss observed for the non-transparent intrusion response with unbuffered intrusion detection corresponds to the loss observed for the first version of location-based intrusion response presented in Chapter 3. Due to the slight overload situation, the overall loss is slightly increased for the non-transparent intrusion response with buffered intrusion detection. Since packets buffered by intermediate nodes are considered lost if we do not take into account retransmissions, the overall packet loss without retransmissions is considerably higher for the transparent intrusion response schemes.

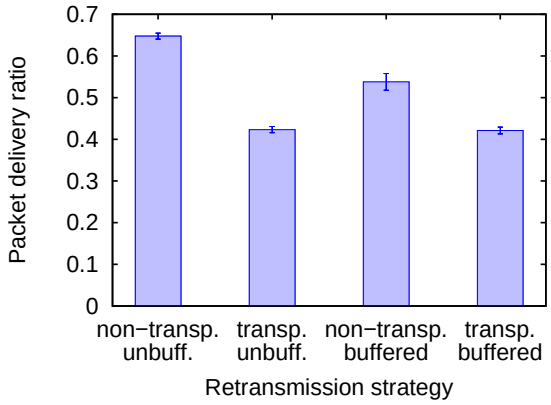
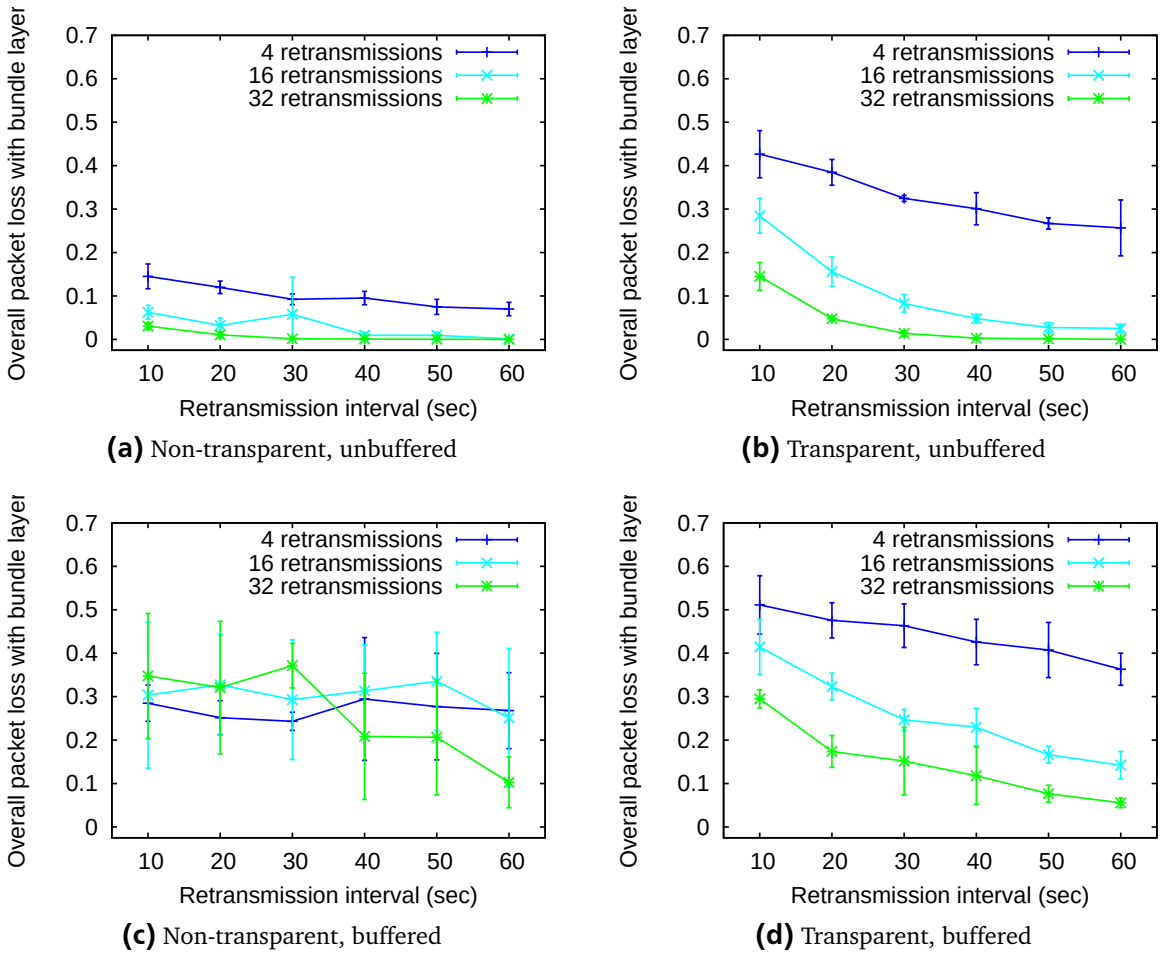


Figure 5.12: Overall delivery ratio without retransmissions of the bundle layer

To verify the correctness of the result, we additionally monitor the packet delivery ratio without retransmissions of the Bundle Layer shown in Figure 5.12. Since packet delivery is monitored indepen-

dently from packet loss and the results for packet loss and packet delivery match, we conclude the correctness of the implementation and of the results presented.

If we take into account the retransmissions of the Bundle Layer, the results change considerably, as shown in Figures 5.13 and 5.14. To interpret the results, we always have to consider both loss and delivery ratios.



(a) Non-transparent, unbuffered

(b) Transparent, unbuffered

(c) Non-transparent, buffered

(d) Transparent, buffered

Figure 5.13: Overall packet loss after retransmissions of the bundle layer

We observe a packet loss reduced to zero for the non-transparent and transparent intrusion response with unbuffered intrusion detection if the maximum number of retransmissions and the time between consecutive retransmissions is chosen appropriately, as shown in Figures 5.13a and 5.13b. Yet, if we consider the corresponding packet delivery ratios shown in Figures 5.14a and 5.14b, we clearly observe delivery ratios below 100%. We put this down to packets still being buffered by the Bundle Layer at the end of the simulated time. These packet were neither dropped nor delivered, thus affecting neither packet loss nor delivery ratios.

On the other hand, if we compare loss and delivery ratios for the non-transparent and transparent intrusion response with buffered intrusion detection, we find that loss and delivery ratios match, as shown in Figures 5.13c and 5.13d and in Figures 5.14c and 5.14d, respectively. Thus, all packets are either delivered or finally dropped by the Bundle Layer at the end of the simulated time.

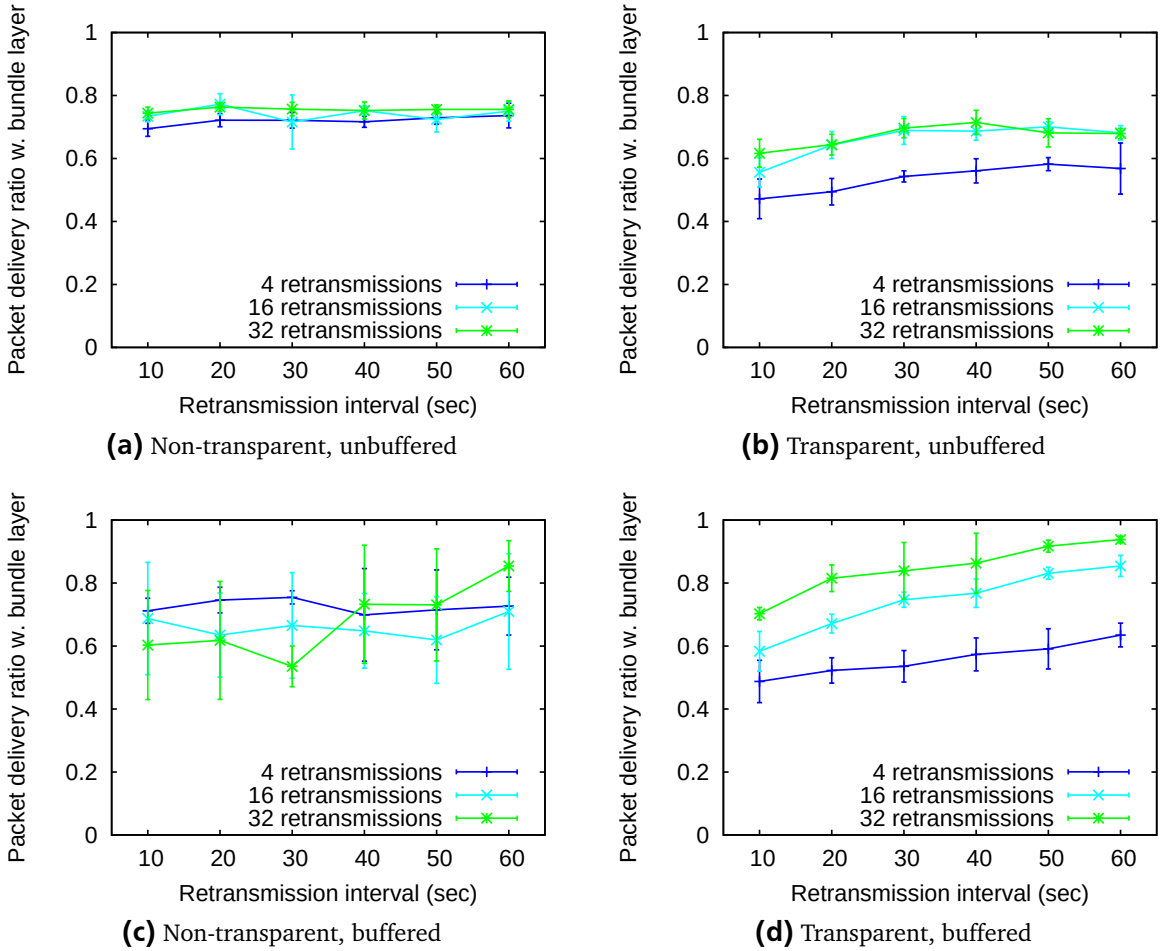
The interpretation above is supported by the transmission delay observed if retransmissions of the Bundle Layer are included in the calculation, as shown in Figure 5.19. We see that the average transmission delay is considerably higher if a buffered intrusion detection system is used.

Comparing the loss observed for a low number of retransmissions and a short time between consecutive retransmissions, we observe higher values for the transparent than for the non-transparent

approaches. This is caused by the fact that more packets have to be handled by the Bundle Layer when a transparent intrusion response is used. Thus, more packets are finally dropped by the Bundle Layer if the maximum number of retransmissions per packet is exceeded.

The unsteady trend and the large confidence intervals we observe for both packet loss and delivery ratios shown in Figures 5.13c and 5.14c can be explained by the slight overload situation that is caused by this combination.

Altogether, we have to emphasize that we are able to increase the packet delivery ratio to more than 90% in case of the transparent intrusion response with a buffered intrusion detection, if appropriate values for the number of retransmissions and the time between consecutive retransmissions are chosen. Recall that the packet delivery ratio in a network without intrusion detection and response is less than 10% for an equal number of misbehaving nodes.



(a) Non-transparent, unbuffered

(b) Transparent, unbuffered

(c) Non-transparent, buffered

(d) Transparent, buffered

Figure 5.14: Overall delivery ratio including retransmissions of the bundle layer

Route Lengths

The average route length observed for all combinations of transparent/non-transparent intrusion response and buffered/unbuffered intrusion detection is shown in Figure 5.15. The baseline given by non-transparent intrusion response combined with unbuffered intrusion detection matches the results presented in Chapter 3 for the first version of location-based intrusion response.

Routes are marginally longer for the transparent intrusion response schemes, which is due to the increased false positive rates of the intrusion detection system, we observe in these scenarios. Increased false positives lead to an increased number of quarantined areas that have to be bypassed by routes. Thus, we obtain slightly longer routes for a transparent intrusion response.

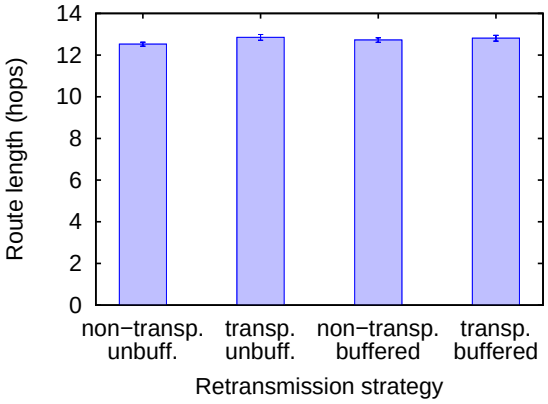


Figure 5.15: Route length

Network Delays

The transmission delays from sending application to receiving application without retransmissions of the bundle layer are shown in Figure 5.16.

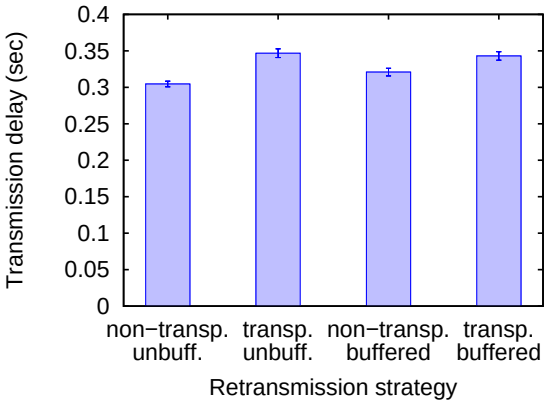


Figure 5.16: Transmission delay from application to application without retransmissions of the bundle layer

Both factors of the transmission delay, the routing delay on the source node and the propagation delay of the network without retransmissions, are shown in Figures 5.17 and 5.18. The baseline given by the delay for non-transparent intrusion response with unbuffered intrusion detection is slightly lower than the delay observed in the first version of location-based intrusion response discussed in Section 3.2.9. This is due to the traffic load which is reduced from 20 to 10 streams running in parallel. If less packets are in flight at a point in time, the queuing delay on each node of a multi-hop route decreases. Thus,

both data packets and routing messages experience a lower delay, resulting in both a decreased routing delay and a decreased propagation delay.

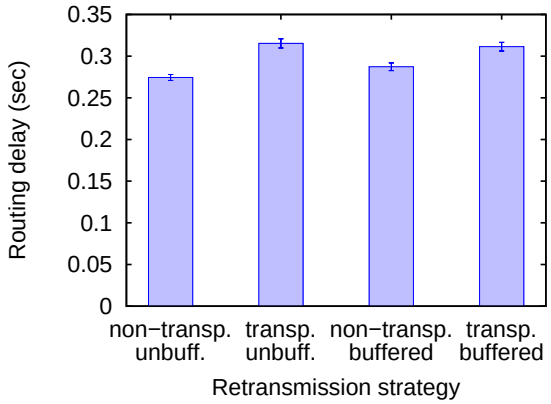


Figure 5.17: Routing delay on source node without retransmissions of the bundle layer

Compared to the baseline, the delays observed for non-transparent intrusion response with buffered intrusion detection are higher which is due to the slight overload situation for this combination.

The delays experienced for transparent intrusion response with both unbuffered and buffered intrusion detection are also increased compared to the baseline. We put this down to an increased network load which is caused by the fact that the number of traffic sources is increased when using transparent intrusion response due the increased false positive rate of the intrusion detection system. The effect starts at the first node that detects a misbehaving node. Since the detecting node does not send a route error, but buffers (thus, not forwards) arriving packets, the detecting node may be falsely classified as black hole, as discussed earlier in this section. With each false positive, a new node buffers packets and, thus, becomes a traffic source as soon as it is not quarantined any more.

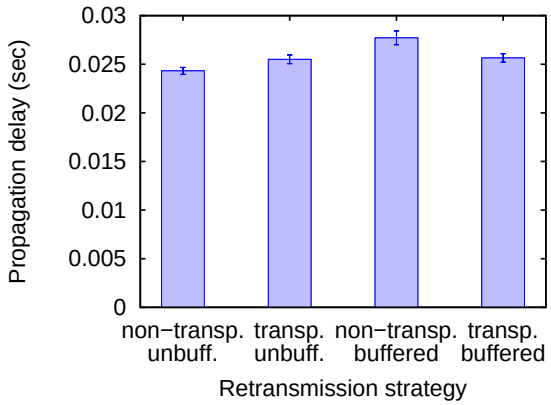
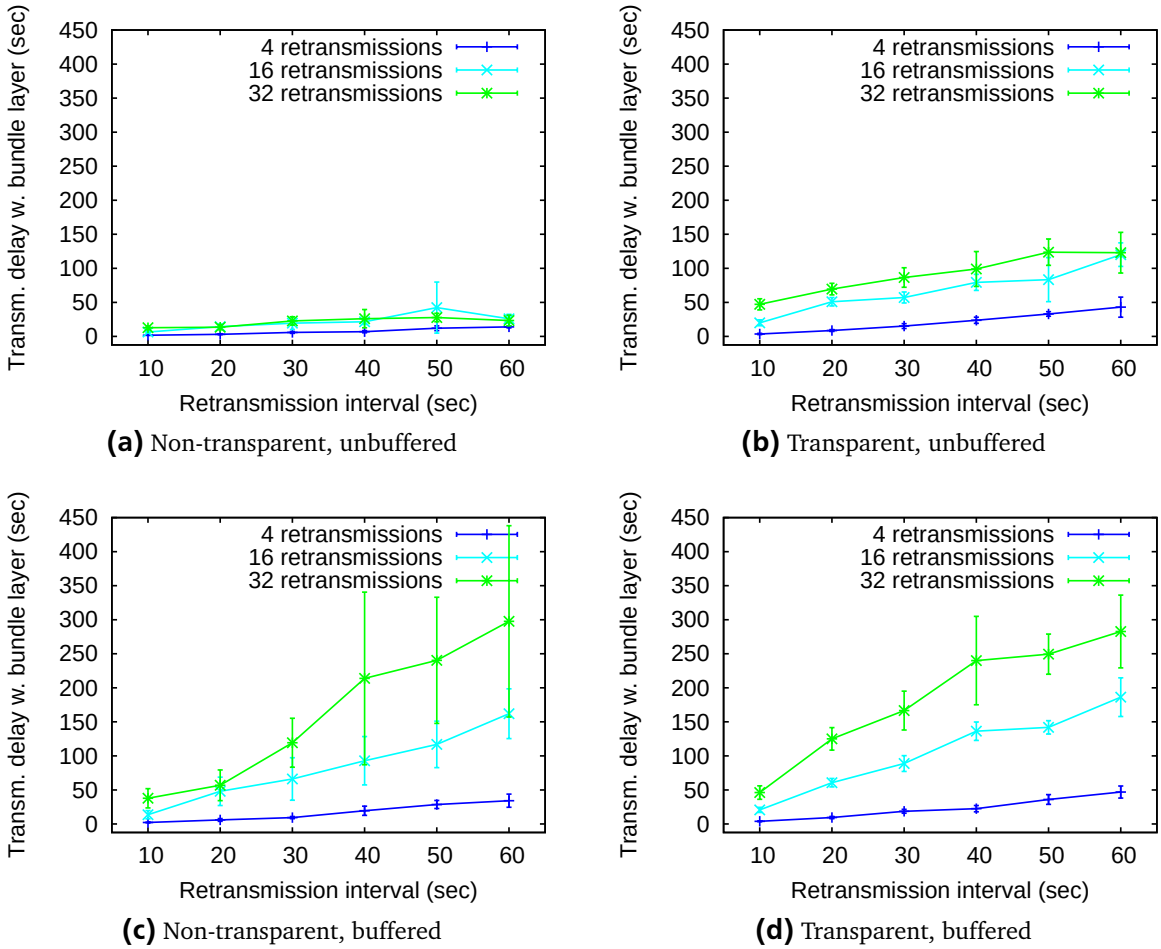


Figure 5.18: Propagation delay from source to destination without retransmissions of the bundle layer

The transmission delay from sending application to receiving application including retransmissions of the bundle layer is shown in Figure 5.19. The lowest delay is observed for non-transparent intrusion response with unbuffered intrusion detection. Compared to all other combinations, in this case, the lowest number of packets has to buffered and retransmitted. Only packets sent to quarantined benign nodes until the source receives the route error are affected. Following this reasoning, the delays observed increase along with the number of packets that are handled by the Bundle Layer. The maximum delays, up to about five minutes, are observed for a transparent intrusion response with buffered intrusion detection, where the highest number of packets is handled by the Bundle Layer. The large confidence

intervals we observe for non-transparent intrusion response with buffered intrusion detection are due to the slight overload situation caused by this combination.



(a) Non-transparent, unbuffered

(b) Transparent, unbuffered

(c) Non-transparent, buffered

(d) Transparent, buffered

Figure 5.19: Transmission delay from application to application including retransmissions of the bundle layer

5.3 Conclusion - Delay Tolerance

In this chapter, we discussed how artificial transmission delays can be used to support the location-based intrusion response in mobile ad hoc networks. By delaying transmissions and retransmissions appropriately, we enable communication of quarantined benign nodes. The evaluation of the different strategies for buffering and retransmitting packets showed that we are able to nearly recover the performance of a clean network with respect to the ratio of packets delivered. The price we have to pay for this is a considerably increased transmission delay. Thus, the approach can be deployed for delay tolerant applications such as e-mail or short message services but is infeasible for real-time communication.

6 Basics and Related Work on Security in Peer-to-Peer Systems

In this chapter, we shortly introduce building blocks of peer-to-peer systems. For a more comprehensive introduction, we refer to [102]. We present basic and recent related work on security that influenced our research. We focus on related work on threshold cryptography and applications of threshold cryptography in peer-to-peer systems. For a broader overview on security in peer-to-peer systems, we refer to [110, 102].

6.1 General Network Model

We consider a spontaneously established peer-to-peer network without central, trusted instances and security policies. We assume that users are involved in security related decisions such as authentication and access control. We focus on finding a set of users that are authorized to take part in security-related decisions and on mapping the decision process to the network.

We assume that the delay introduced by involving users in decision is in orders of magnitude higher than delays introduced by the network or by means of cryptography. We, thus, neglect network and computation delays. We demand that a decision process should be completed within one round. That is, after one request-reply cycle, a decision should be found.

Regarding the peer-to-peer system, we demand a reliable and deterministic lookup process, that is, assuming a lossless network, a lookup for a particular service or user available in the peer-to-peer system should be successful as soon as the service or the user is available in the system.

6.2 Foundations

In contrast to client/server systems, where the roles of service provider and service consumer are separated strictly, peers in a peer-to-peer system act as both service provider and service consumer. This way, the application layer of a communication network can be established spontaneously, without the need of setting up central instances. A core challenge in peer-to-peer systems (and in communication networks in general) is to lookup a requested service, that is, to determine the network address at which the service can be reached. In client/server systems, this task is performed by centralized service directories such as the domain name system or web search engines. During the evolution peer-to-peer systems have undergone in the last years, several approaches for service lookup were proposed. For our needs, these approaches can be used to categorize peer-to-peer systems into centralized, pure, hybrid, and structured peer-to-peer systems.

In centralized peer-to-peer systems, a central instance is responsible for service lookup. A service provider registers the service at the central instance. A service consumer can perform a service lookup by contacting the central instance. In large-scale, Internet-based peer-to-peer systems, this approach introduces challenges regarding scalability due to a potentially high traffic and computing load of the central instance. In our scenario, where the peer-to-peer system is established on top of a mobile ad hoc network, the central instance may become disconnected, thus rendering service lookup impossible.

In pure peer-to-peer systems, service lookup is, in general, performed by broadcasting a lookup request to all peers in the overlay. This approach may cause overload situations and result in an unreliable, non-deterministic lookup process [102].

To overcome the limitations of pure peer-to-peer approaches, hybrid systems introduce a two-tier architecture for service lookup, thus combining advantages of centralized and pure peer-to-peer systems. For this, a a set of dedicated peers, which usually have high resources, is responsible for service lookup. When joining a peer-to-peer system, each 'normal' peer connects to a dedicated peer at which it registers

available services. Lookup requests are sent to the dedicated peer which contacts other dedicated peers if the service requested is not registered at itself. While the traffic and computing load caused by lookup requests can be reduced this way, still, the lookup process remains non-deterministic. Due to load limitation strategies, a lookup request may not reach the dedicated peer, at which a service is registered.

In structured peer-to-peer systems, each peer is responsible for the lookup information of a subset of the services available in the system. For this, peers and services share a common address space. A peer offering a service registers the service at a peer with an address closest to the service description. Taking the high-level analogy of a phone book, services with descriptions from 'aardvark' to 'azure' are registered at Peer 'a', et cetera. Lookup requests are directed to the peer with the address closest to the service description, resulting in a deterministic lookup process. That is, a lookup request will be successful as soon as the requested service is available in the peer-to-peer system. Note that, in our scenario, this only holds if the mobile ad hoc network is not partitioned, that is, if the peers involved in service lookup are not disconnected from the network.

Recently, for example in [82, 18, 57], architectures for structured peer-to-peer systems that are tailored to the challenging conditions of mobile ad hoc networks were proposed. Yet, we do not focus on the challenge of implementing a structured peer-to-peer system on top of a mobile ad hoc network in the following. Further, the models and tools developed for user-based cooperative decisions in peer-to-peer systems do not depend on a particular implementation of a peer-to-peer overlay. Also, due to the lack of a mobile peer-to-peer testbed of appropriate size, the evaluation presented in the next chapter is based on wired, static testbeds. Without loss of generality, we thus decided to use the Pastry structured peer-to-peer overlay [88] described below as basis for our experiments.

6.2.1 Pastry

Pastry [88] is a representative of structured peer-to-peer systems. Service lookup in Pastry is based on a circular address space, with n bit addresses. Commonly, $n = 128$ is chosen. Both peers and services are addressed in this address space. Service addresses are maintained by the peer with the address closest to the particular service address.

To route lookup requests to the responsible peer, each node maintains a prefix-based routing table which consists of $n \cdot 2^{-b}$ rows and 2^b columns, where 2^b is the base in which addresses are expressed numerically. Commonly, b=4 is chosen. In Row r of the routing table, known peers with an address matching the first r bits of the node maintaining the routing table are stored. This results in a routing table that is dense in the logical neighborhood of a peer and sparse for distant parts of the address space, as shown schematically in Figure 6.1. The metric used for populating the routing table takes into account the physical distance of peers, determined, for example, by the round trip time.

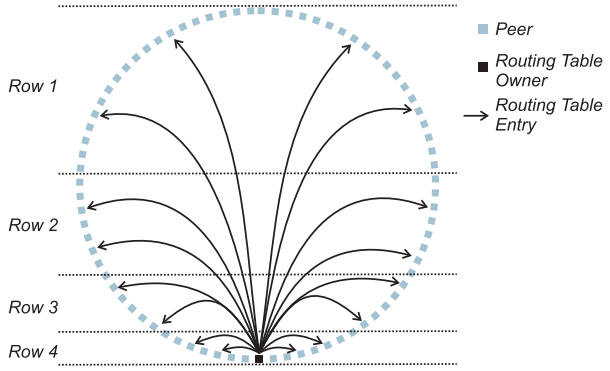


Figure 6.1: Schematic representation of Pastry's routing table for $b = 2$. Note that the number of rows does not match the specification for reasons of presentation.

In addition to the routing table, nodes maintain a neighborhood set and a leafset. The neighborhood set contains peers that are close with respect to the routing metric, that is, physically close peers. The leafset contains peers that are close with respect to peer addresses, that is, logically close peers. The leafset typically consists of 24 peers. Both neighborhood set and leafset are used to improve routing performance in that a lookup request is preferably forwarded to nodes logically closer to the requested service address based on either the neighborhood set or the leafset. The routing table is used if no appropriate node is found in the neighborhood set or the leafset.

6.2.2 Scribe

Scribe [16] enables application layer multicast communication in Pastry peer-to-peer systems. Peers can flexibly create, join, and leave multicast groups. For each group, a common multicast tree is established and used to forward multicast messages. In analogy to Pastry's addressing scheme, the peer with the address closest to the name of the multicast group acts as entry point, that is, as root of the multicast tree. To send a multicast message to a group, the message is first sent to the entry point and, from there, forwarded along the multicast tree.

A peer that wants to join a multicast group sends a corresponding join message to the entry point of the group using the routing mechanism of Pastry. Each peer that forwards the join message adds the peer from which it received the message to the set of its children in the corresponding multicast tree if it is not included already. This process is stopped as soon as a peer, from which a join message is received, is already included in the set of child peers or if the join message reaches the entry point.

If a peer leaves a tree, a leave message is sent upwards the tree, pruning all branches that are not required anymore, that is, all branches that were needed to reach the leaving peer only.

Tree maintenance is performed in a proactive way. For this, heartbeat messages are sent periodically through the tree. Nodes that do not receive heartbeat messages for a tree they are part of send a new join message, thus repairing the tree. To handle failing entry points, their functionality is replicated to logical neighbors. This way, the routing mechanism of Pastry takes care of reaching the new entry point in case of a failure.

6.3 Related Work on Threshold Signatures

The mathematical foundation behind achieving joint decisions can be provided by threshold signatures. The idea of threshold signatures was first described in [95]. The mode of operation is based on a (private) signature key K that is split up into a set of partial keys K_i with $i \in \{1, ..., n\}$. For this, a polynomial $q(x)$ is chosen such that $K = q(0)$ applies. The partial keys then correspond to values of $q(x)$ at specific indexes, that is $K_i = q(i)$ where $i \neq 0$. The application scenario considered in [95] includes a trusted signing device that collects partial keys whenever a signature has to be produced. By Lagrange interpolation, the trusted device can compute the full signature key and produce a valid signature if enough partial keys are collected. Thereby, 'enough' is defined by the degree of the chosen polynomial $q(x)$.

The availability of a trusted signing device can not be assumed in our highly dynamic peer-to-peer scenario. For such environments threshold signature schemes which do not interpolate keys but signatures were proposed. Every peer P_i that is authorized to take part in a joint decision receives a partial key K_i. If, for example, a certificate C which allows for the access to restricted services of a closed user group has to be signed, each peer that holds a partial key K_i may issue a certificate with a partial signature S_i. By Lagrange interpolation, the peer that has requested access can interpolate the full signature S from the partial signatures if a sufficient number of partial signatures is received (that is, if sufficient authorized peers decided to grant access).

For our needs, we have to differentiate between interactive and non-interactive threshold cryptography schemes. Unlike non-interactive approaches, interactive ones require communication (in possibly

multiple rounds) between the peers that want to issue partially signed certificates. We illustrate this point using a simplified example of threshold RSA signatures. A 'standard' RSA signature $S(m)$ of a message m is produced by calculating

$$S(m) = m^d \mod N$$

where d is the private exponent and N is the RSA modulus. When threshold RSA signatures are used, d is split up in a set of partial exponents d_i. In interactive approaches, the partial signature $S_i(m)$ of Peer P_i is calculated as

$$S_i(m) = m^{d_i \cdot L_i} \mod N$$

where L_i is the Lagrange factor of P_i with

$$L_i = \prod_{j=0, j\neq i}^{k} \frac{-x_j}{x_i - xj}$$

Subsequently, each Peer P_i needs to know the indexes of all other $k-1$ peers $P_{j\neq i}$ that provide partial signatures so as to be able to calculate L_i. In order to obtain this knowledge, the peers involved have to exchange their indexes. This further implies that a full signature cannot be interpolated if a partial signature is lost or if a peer does not provide its partial signature within a reasonable amount of time. In this case, a new request has to be sent since the partial signatures are only valid in the signature group for which the Lagrange factors have been calculated.

Due to the restrictions introduced by involving users in the decision process that may not be able to provide a decision in a reasonable time, interactive threshold cryptography approaches are not suitable for our application scenario. Appropriate non-interactive approaches are presented, for example, in [96, 21]. Here, the generation of a partial signature has been modified such that knowledge on the indexes of the other peers contributing partial signatures is not required. This property is paid for by the loss of other functionality offered by interactive threshold signature schemes. As an example, in [96] a trusted dealer that generates and distributes the partial keys is assumed, whereas other threshold signature schemes allow for the distributed calculation of the keyshares. In our application scenario this does not pose a problem. The distribution of the partial keys can be performed a priori, within a secure environment (for example, by cell phone operators) where a trusted dealer is available. A more detailed discussion of this point is beyond the scope of this thesis. For our scenario, gaining the advantage of non-interactivity is worth the price of loosing other features.

A signature S that was produced by threshold cryptography does not reveal anything about the partial keys K_i that were used. Thus, no information about the peers that were involved in a joint decision can be deduced from a signature. Also a restriction to a subset of the shareholders is not possible. For some scenarios, it may be necessary to identify single peers that were involved in a joint decision or to define additional restrictions to decide which peers are allowed to take part in a joint decision. In this case, multisignatures can be used instead of threshold cryptography as the basic technique. Here, every authorized peer owns a (full) signature key which is uniquely bound to its identity. A simple multisignature scheme would be to form a signature by the concatenation of the identities and the signatures of the signers. However, this scheme would cause the size, as well as the computing time required for the verification of a signature to grow linearly with the number of signers. Since the bandwidth available and the computing power of devices may be restricted in our scenario, this simple multisignature scheme does not meet our requirements. More complex schemes produce compressed signatures without the loss of information. Here, again we have to differentiate between interactive and non-interactive schemes.

In analogy to the reasons provided in the context of threshold cryptography, interactive multisignature schemes are not appropriate for our application scenario. A suitable non-interactive scheme has been proposed, for example, in [11].

A drawback of applying public key cryptography in resource constraint networks is that certificates have to be exchanged to learn public keys of communication partners. This poses a data overhead for a bandwidth constraint network. To avoid this data overhead, the cryptosystem can be made self-contained, meaning that each node is preloaded with the public keys of all other nodes. This, on the other hand, may cause a significant memory overhead for resource constraint devices. A mechanism for self-contained key management in resource constraint ad hoc networks is proposed in [34]. To reduce memory overhead, nodes are identified by a particular set of private keys, thus decreasing the overall number of public/private keypairs required.

6.4 Related Work on Applying Threshold Cryptography in Peer-to-Peer Systems

In [73], the authors compare different threshold signature schemes with respect to their performance in controlling access to closed user groups within peer-to-peer systems. Performance is measured in terms of basic operation costs, that is, the time needed to produce partial signatures, and join time, that is, the amount of time a new peer needs to join a closed user group. User interactions, as required in our application scenario, are not considered. Non-interactive signature schemes which we assume to be the most promising variant of threshold signatures for our application scenario are not part of this evaluation. The analysis of [73] is extended to fit additional signature schemes in [90].

In [62] and [63], a protocol for access control in mobile ad hoc networks using interactive threshold cryptography is proposed. The protocol is based upon localized joint authentication within the geographical neighborhood of a node. The performance of the approach proposed is evaluated (amongst other metrics) regarding the probability that an authentication request is successful, that is, whether a sufficient number of partially signed certificates has been received to be able to interpolate a full signature. Stochastic models for this success probability are not proposed. The protocol is proven to be susceptible to attacks in [47].

A non-interactive mechanism for access control to mobile ad hoc networks is proposed in [91] and [116]. In contrast to [96], which we identified as a threshold signature scheme that is applicable in our application scenario, the protocol proposed in [91] and [116] is not based on a cryptographic key that is shared among multiple parties, but on bivariate polynomials that can be used to establish pairwise shared secret keys. A performance evaluation comparable to [73] and [90] has been carried out. User interactions are not considered. Because 'standard' signed certificates as required in our scenario cannot be produced with this approach, we did not choose [91] as a possible cryptographic mechanism.

6.5 Related Work on Analytical Models for Peer-to-Peer Systems

A large variety of related work on analytical models for peer-to-peer systems exists. Mostly, the peer-to-peer system itself is considered by modeling metrics such as network load, computational load, and quality of service parameters.

In [51], the authors develop a model for controlling the node degree of peers in a pure peer-to-peer system. The model aims at describing the trade-off between traffic overhead experienced by individual nodes that is caused by lookup operations and content availability, that is, the probability that a service available in the peer-to-peer system can be found. By adjusting the node degree, both factors can be balanced.

Effects of up and downscaling peer-to-peer filesharing systems are described analytically in [61]. Metrics like content availability, lookup delays, and transmission delays subject to the size of the peer-to-peer system are considered. The model is tailored to hybrid peer-to-peer systems.

Scalability and performance of structured peer-to-peer systems are modeled in [111]. The focus is set on evaluating the benefits and drawbacks of building a structured peer-to-peer system upon few highly available nodes with rich resources or upon many nodes with potentially restricted resources that join and leave the peer-to-peer system frequently, thus causing a large amount of maintenance traffic.

The effects of churn, that is, peers frequently joining and leaving, on a structured peer-to-peer system are further studied in [55]. The focus is set on calculating the probability of a partitioned overlay.

The model developed in [58] describes the effects of caches for peer-to-peer traffic installed at service providers. The focus is set on modeling the traffic that crosses domains of Internet service providers caused by peer-to-peer systems.

A model describing the computational load and the network load of peer-to-peer streaming systems is proposed in [17]. By taking into account the topology of the underlay network, the model is tailored to proximity-based streaming systems. That is, the model describes the scalability of streaming solutions that select peers from which to download based on the physical proximity of the peers in the underlay network.

In [120], a model for controlling the scheduling of data chunks in peer-to-peer streaming systems is proposed. The goal of the model is to provide a means for describing the trade-off between spreading data chunks within the peer-to-peer system to increase scalability and delivering content to a particular peer to increase quality of experience in terms of playback smoothness.

An analysis of different variants of authentication protocols with respect to lower bounds of message overhead generated and communication rounds required is performed in [26]. Protocols based on threshold signatures are not considered. Thus, stochastic models for different interaction schemes in a joint process are not proposed.

To the best of our knowledge, stochastic models that describe different interaction schemes for user-based joint decisions based on non-interactive threshold signature mechanisms, which are the focus of our work, were not proposed.

7 User-based Cooperative Decisions in Peer-to-Peer Systems

In this chapter, we present the concept of user-based cooperative decisions. Figure 7.1 shows a schematic representation of a cooperative decision process based on threshold cryptography using the example of access control to a restricted resource. In Step 1, the user that requires access to the resource sends a corresponding request to a set of peers that are authorized for deciding on security-related requests, that is, that hold keyshares. In Step 2, each peer requested involves its user in the decision process. The request is presented in a human-readable form to the user, who may, if he/she is not otherwise engaged, agree to granting access. In this case, the peer sends a certificate partially signed with the peer's keyshare back to the requesting peer. In Step 3, the partially signed certificates can be combined to a fully signed certificate if a sufficiently large number is collected. This way, user-based cooperative decision offer an alternative to central, trusted instances for building the fundament of security services like authentication and access control.

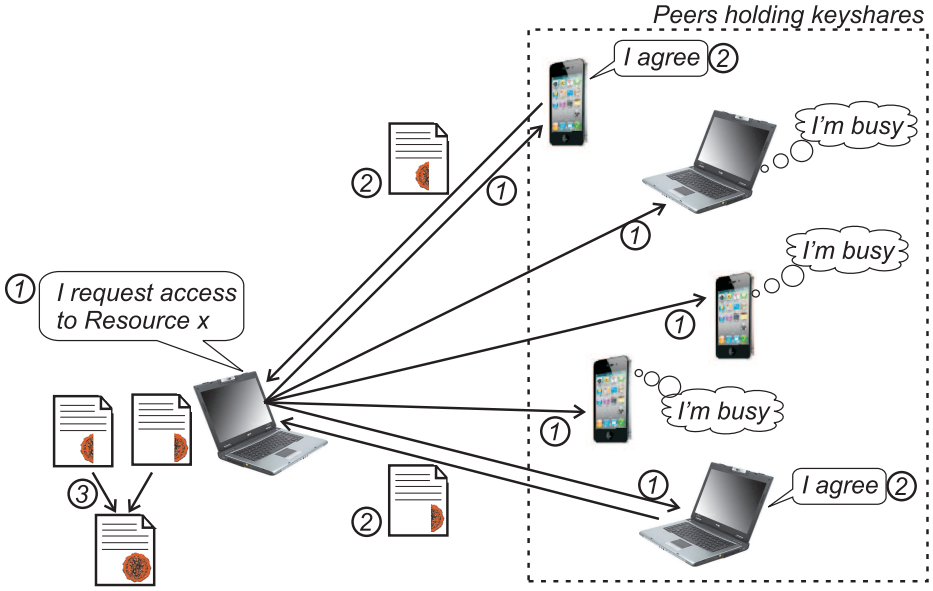


Figure 7.1: Schematic representation of a cooperative decision based on threshold cryptography

In the following, we first discuss schemes for the interaction of a peer that request a security-related decision with the peers that are involved in the decision process. Based on this, we develop stochastic models describing the success probability of of a decision process subject to the particular interaction scheme. We consider a decision process successful, if the requesting peer receives a sufficiently large number of partially signed certificates to be able to compute a fully signed certificate. We verify the models developed with a prototypic implementation of user-based cooperative decisions in the PlanetLab [80] and the G-Lab [19] testbeds.

7.1 Architecture

As shown in Figure 7.1, a peer that issues a security-related request has to send this request to a set of peers that hold keyshares and, thus, are authorized to take part in a decision process. The strategy for disseminating the request within the peer-to-peer system directly affects the number of users requested and the probability of receiving enough partially signed certificates to be able to interpolate a full signature. In the following, we describe different interaction schemes between requesting peers and peers that (potentially) contribute to the decision process. We take into account different levels of knowledge

about which peers hold keyshares. We shortly explain how the interaction schemes can be realized in a Pastry peer-to-peer overlay, which will be the basis for the testbed evaluation. However, the interaction schemes proposed in the following are independent from the particular peer-to-peer system and hold for structured as well as for unstructured or hybrid systems.

7.1.1 Interaction Scheme for Known Shareholders

If we assume knowledge on the distribution of keyshares and on whether the corresponding authorized users are available for contributing to a decision process, we can harness this knowledge to coordinate the dissemination of requests.

This can be done, for example, by a peer that acts as mediator for the decision process. The mediator keeps track of the status of authorized peers/users and accepts and relays requests appropriately, as shown in Figure 7.2.

As an alternative, the mediation can be implemented based on Scribe multicast groups. Peers that hold keyshares, with users ready to contribute to a decision process, subscribe to a corresponding multicast group. Requests are sent to this group along with the requested number of partial signatures.

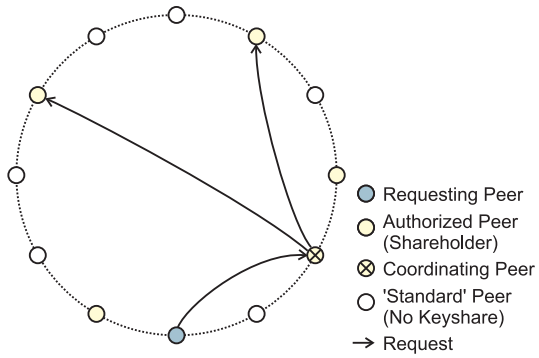


Figure 7.2: Schematic representation of the interaction scheme with known shareholders

7.1.2 Interaction Scheme for Unknown Shareholders

If no knowledge on which peers hold keyshares is available, broadcasting the request within the entire peer-to-peer system would be a straightforward interaction scheme. In this case, all users holding keyshares are contacted. Thus, in case the decision is positive, a broadcast would result in the highest probability for a decision request to be successful (that is, for receiving enough partially signed certificates to be able to interpolate a fully signed certificate). Yet, due to the high number of users involved that is caused by a broadcast in the peer-to-peer overlay, the applicability of a broadcast is limited in our scenario. Instead, a multicast approach is reasonable. We assume that the multicast is initiated by the requesting peer and that the requesting peer has no knowledge about which peers hold keyshares.

One way to realize a multicast is to send requests to a set of randomly selected peer IDs, as shown in Figure 7.3. Regarding the implementation in a Pastry peer-to-peer system, if an ID is not used, Pastry will route the request to the peer with the ID closest to the one selected.

Another multicast approach that can be applied if the shareholders are unknown to the requesting peer is based on gossiping the request among neighbor peers, as shown in Figure 7.4. In this case, the definition of neighbors depends on the metrics of the particular peer-to-peer system deployed. For Pastry, we implemented a forwarding based on the leafset. A requesting peer sends the request to the peers in its leafset. Each peer that receives the request again forwards the request to the logically most distant peer in its leafset. To limit the propagation of the request, a counter indicating the required number of partial signatures is contained in the request message. The peer that initiates the request sets the

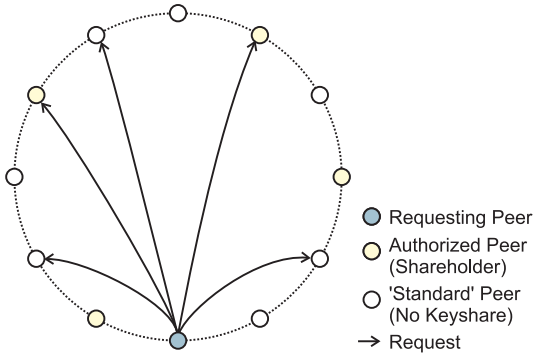


Figure 7.3: Schematic representation of the interaction scheme for unknown shareholders. Requests are sent to randomly selected peer IDs.

counter to the total number of partial signatures required divided by the number of peers in its leafset. Each shareholder that receives the request decreases the counter appropriately. Request messages with an expired counter are dropped.

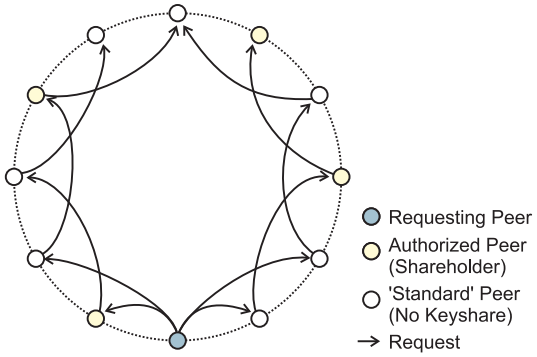


Figure 7.4: Schematic representation of the interaction scheme for unknown shareholders. Requests are gossiped to neighbor peers.

7.2 Stochastic Analysis of the Decision Process

In this section, we develop stochastic models that describe the success probability p_{succ} of a cooperative decision process subject to the interaction schemes described in the previous section. A request is considered successful if a sufficient number of partially signed certificates is received such that a fully signed certificate can be interpolated. The number of partially signed certificates received is sufficient if it is greater or equal to the threshold n_{thres} that is defined by the threshold cryptography scheme deployed. Table 7.1 provides an overview on the notation we use in the following.

We provide graphical examples of the models using a scenario that consists of $n_{total} = 100$ peers of which $n_{keys} = 15$ hold keyshares. The probability p_{rep} that a peer holding a keyshare contributes to a decision process in a reasonable time is assumed to be 95%. With these values given, we show how n_{thres} and n_{req} affect the success probability p_{succ}. Please note that the values are chosen for visualization purposes such that the system operates in reasonable bounds. The stochastic models are independent from this particular instantiation.

7.2.1 Model of the Interaction Scheme for Known Shareholders

For the interaction scheme with known shareholders, the probability $p(n_{rep})$ of receiving a certain number n_{rep} of replies to a request can be modeled as a binomial random variable. A binomial random

Table 7.1: Notations of formulae

n_{thres}	Number of partially signed certificates required to compute a valid certificate
p_{rep}	Probability with which a single peer answers a request
n_{total}	Total number of peers in the peer-to-peer system
n_{keys}	Number of peers holding keyshares
n_{req}	Number of peers to which a request is sent
n_{rep}	Number of replies received to one request
$p(n_{rep})$	Probability for receiving exactly n_{rep} replies to one request
p_{succ}	Probability for receiving a sufficient (subject to n_{thres}) number of replies to one request

variable describes the outcome of repeated Bernoulli experiments each of which has a certain probability of being successful. In our case, the probability of being successful is given by p_{rep} and the number of repetitions is the number n_{req} of peers to which a request is sent. This results in

$$p(n_{rep}) = \binom{n_{req}}{n_{rep}} p_{rep}^{n_{rep}} (1 - p_{rep})^{n_{req} - n_{rep}}$$

For a request to be successful it is not important to receive a certain number of replies, but to receive a sufficiently large number to enable computation of a fully signed certificate. That is, a request is successful if the number n_{rep} of replies received is greater than or equal to n_{thres}. The success probability $p_{succ}(n_{thres})$ subject to n_{thres} thus can be described as the sum of the probabilities of receiving a certain number n_{rep} of replies starting from n_{thres}. The upper bound of the sum is given by the number n_{mult} of peers to which a request is sent. We get

$$p_{succ}(n_{thres}) = p(n_{rep} \geq n_{thres}) = \sum_{n_{rep}=n_{thres}}^{n_{req}} p(n_{rep})$$

The success probability subject to n_{thres} (note that for reasons of presentation, the axis is inverted) and n_{req} is shown in Figure 7.5. The success probability increases along with the number of users requested and decreases if n_{thres} increases.

7.2.2 Model of the Interaction Scheme for Unknown Shareholders

For the interaction scheme for unknown shareholders, we assume a random limitation of requested peers subject to the distribution of keyshares and subject to the status of peers. That is, the interaction scheme does not consider whether a peer to which a request is sent is authorized to contribute to security-related decisions or able to answer within an acceptable time frame. This random restriction can be modeled by a hypergeometric random variable. In our case, the hypergeometric variable describes the intersection of the set of peers to which a request is sent and the set of peers that would potentially contribute to a decision process. Thus, for the probability $p(n_{rep})$ of receiving a certain number n_{rep} of replies, we get

$$p(n_{rep}) = \frac{\binom{n_{keys} \cdot p_{rep}}{n_{rep}} \binom{n_{total} - (n_{keys} \cdot p_{rep})}{n_{req} - n_{rep}}}{\binom{n_{total}}{n_{req}}}$$

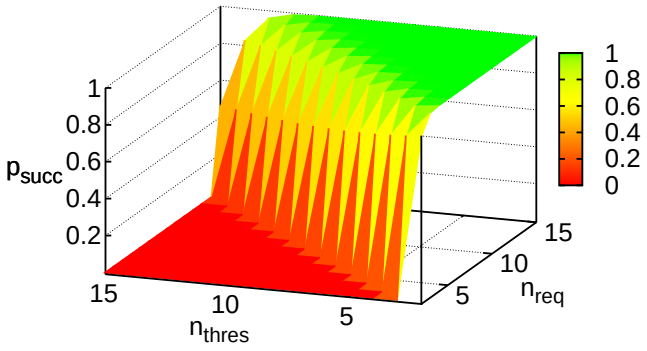


Figure 7.5: Success probability $p_{succ}(n_{thres})$ of the interaction scheme with known shareholders.

As for the interaction with known shareholders, a request is successful if the number n_{rep} of replies received is greater than or equal to n_{thres}. The success probability $p_{succ}(n_{thres})$, thus, is again the sum of the probabilities of receiving a certain number n_{rep} of replies starting from n_{thres}. The upper bound of the sum is given by the number n_{req} of peers to which a request is sent. We obtain

$$p_{succ}(n_{thres}) = p(n_{rep} \geq n_{thres}) = \sum_{n_{rep}=n_{thres}}^{n_{req}} p(n_{rep})$$

The resulting success probability subject to n_{thres} (axis inverted) and n_{req} is shown in Figure 7.6. As for the interaction scheme with known shareholders, the success probability decreases if n_{thres} increases and increases along with the number of users requested.

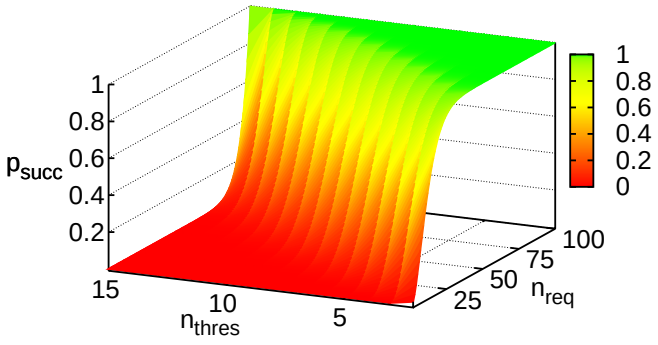


Figure 7.6: Success probability $p_{succ}(n_{thres})$ of the interaction scheme with unknown shareholders

If n_{req} is sufficiently small with respect to n_{total}, the hypergeometric random variable can be approximated by a binomial random variable. Since our goal is to send a request to the least number of peers possible, the binomial approximation to the hypergeometric random variable is applicable for our needs. Figure 7.7 shows the resulting approximation error of the success probability p_{succ}. The approximation error is low for reasonable values of n_{thres} and n_{req} subject to n_{total}. Therefore, we may represent the interaction scheme for unknown shareholders by a binomial random variable. This allows us to unify the models for both interaction schemes.

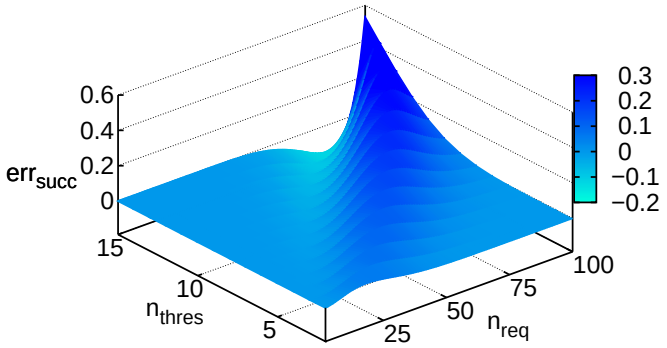


Figure 7.7: Approximation error if the interaction scheme for unknown shareholders is described by a binomial random variable

7.2.3 Closed-Form Representation

We now derive a formula that provides a lower bound for the success probability p_{succ}. Since we can describe both interaction schemes based on a binomial random variable with different parameters, the resulting formula is applicable for all interaction schemes discussed. With this, we aim at being able to adjust the system parameters without knowing the exact network conditions as required for the models presented previously. Since the Chernoff bound is a pessimistic approximation, it may be applicable also if factors such as packet loss and churn are not known exactly but are within reasonable bounds.

Let $p_{fail} = 1 - p_{succ} = p(n_{rep} \leq n_{thres} - 1)$ be the probability that our system fails (that is, not enough partial signatures were received by the requesting peer). We apply Chernoff's bound [87]

$$p(n_{rep} \leq n_{thres} - 1) \leq e^{(-\tau(n_{thres}-1))} M(\tau) \quad \forall \tau < 0$$

to obtain an upper bound for p_{fail}. The moment generating function $M(\tau)$ for n_{rep} is given by

$$M(\tau) = p_{rep} e^{\tau} + (1 - p_{rep})^{n_{mult}}$$

thus,

$$p_{fail} \leq e^{-\tau(n_{thres}-1)} (p_{rep} e^{\tau} + (1 - p_{rep}))^{n_{req}}$$

To obtain an upper bound for p_{fail} (thus, a lower bound for p_{succ}) we have to find the τ that minimizes the right hand side. Let

$$f(\tau) = e^{-\tau(n_{thres}-1)} (p_{rep} e^{\tau} + (1 - p_{rep}))^{n_{req}}$$

thus,

$$\frac{df}{d\tau}(\tau) = e^{-\tau(n_{thres}-1)} n_{req} (p_{rep} e^{\tau} + (1 - p_{rep}))^{n_{req}-1} p_{rep} e^{\tau} - n_{thres} e^{-\tau(n_{thres}-1)} (p_{rep} e^{\tau} + (1 - p_{rep}))^{n_{req}}$$

By setting $\frac{df}{d\tau} = 0$ and resolving to τ, we obtain

$$\tau = \ln\left(\frac{(n_{thres}-1)(1-p_{rep})}{n_{req}p_{rep} - p_{rep}(n_{thres}-1)}\right)$$

Thus, altogether,

$$p_{succ} \geq 1 - \left(\frac{n_{req}p_{rep} - p_{rep}(n_{thres}-1)}{(n_{thres}-1)(1-p_{rep})}\right)^{n_{thres-1}} \left(\frac{(n_{thres}-1)(1-p_{rep})}{n_{req}-n_{thres}+1} + (1-p_{rep})\right)^{n_{req}}$$

The Chernoff bound is applicable as long as $p_{rep} > 0.5$ holds.

7.3 Evaluation

We now present the evaluation of a prototype of user-based cooperative decisions controlled by the models developed previously. As for the evaluation of location-based intrusion response in mobile ad hoc networks, we follow the methodology proposed in [45].

7.3.1 Goals of the Evaluation

The goal of the evaluation is to first provide a proof-of-concept of user-based cooperative decisions. For this, we deploy a prototypic implementation in two testbeds. On this basis, we verify the correctness of the stochastic models describing the interaction schemes presented in Section 7.1 by comparing model predictions and testbed results.

7.3.2 Services of the System

The system under test is a Pastry structured peer-to-peer system as introduced in Section 6.2.1. We implemented the cooperative decision process based on the non-interactive threshold signature scheme presented in [96]. For the interaction of requesting peer and peers that potentially contribute to a decision, the four different schemes for known and unknown shareholders sketched in Sections 7.1.1 and 7.1.2 are available.

7.3.3 Metrics for the Evaluation

The metric we consider for verifying the validity of the stochastic models is the success probability p_{succ} for the different interaction schemes. We determine the deviation of model predictions and results obtained from testbed experiments for both open and closed form models.

7.3.4 Parameters of the System

The parameters of the peer-to-peer system are, as specified in Table 7.1, the number n_{thres} of partially signed certificates required to compute a full signature, the probability p_{rep} with which a single peer answers a request, the total number of peers in the peer-to-peer system n_{total}, the number n_{keys} of peers holding keyshares, and the number n_{req} of peers to which a request is sent. Additionally, inherent parameters of the testbeds such as packet loss, churn, and distribution of peer IDs that are beyond our control affect system behavior.

7.3.5 Selection of the Factors for the Evaluation

For our testbed studies, we vary the total number of peers in the peer-to-peer system n_{total}, the number n_{keys} of keyshares distributed among the peers, and the number n_{thres} of partially signed certificates required to compute a full signature. We choose the parameters such that we obtain variations of few/many peers, few/many keyshares, and high/low threshold.

7.3.6 Evaluation Technique

We validate the models for the different interaction schemes by experiments in the PlanetLab [80] and G-Lab [19] testbeds. PlanetLab is a worldwide testbed currently consisting of about 1000 nodes distributed at about 500 sites that are connected over the Internet. The PlanetLab software offers SSH accounts with a minimal Linux installation. Although minimum requirements regarding hardware and connectivity of nodes are specified, the hardware and the connectivity of nodes in PlanetLab are heterogeneous.

G-Lab, on the other hand, is a German national testbed that currently consists of about 150 nodes at 6 universities. Compared to PlanetLab, the hardware and the connectivity are homogeneous and the overall load of nodes is moderate (by now). G-Lab uses PlanetLab software for resource sharing and access control of its nodes. Thus, our experimental setups can be executed on PlanetLab and on G-Lab without modifications.

7.3.7 Workload of the System

Since our goal is not to evaluate the performance of the Pastry peer-to-peer system or of the testbeds we use to run the experiments, but to verify the correctness of the stochastic models, the workload of the system is low. We only issue one request at a time and wait for replies. This way, we eliminate side-effects of an overloaded peer-to-peer system, which would result in falsified network parameters. Still, we have no influence on the workload of the testbeds, which considerably affects the results as we will discuss in the following.

7.3.8 Experimental Design

To verify the validity of the stochastic models describing the different interaction schemes, we selected the parameters and factors used in the testbed experiments as follows.

Parameters

For the evaluation, the parameters as listed in Table 7.2 are used. In preliminary experiments, we found that the effects caused by varying the probability p_{rep} with which a single peer answers a request are equivalent to the effects caused by varying the number n_{keys} of peers holding keyshares. To keep the number of experiments required to cover the factors used for evaluation in reasonable bounds, we, thus, choose a fixed value for p_{rep} (random numbers following a shifted standard normal distribution are drawn to determine whether a reply should be provided). The same reasoning holds for varying the targeted success probability p_{succ} and the number n_{thres} of partially signed certificates required to compute a valid certificate.

Table 7.2: Parameters as used in the testbed experiments

Targeted success probability	$p_{succ} = 0.95$
Number n_{req} of requests issued	Determined by model to achieve $p_{succ} = 0.95$

Factors

The factors we used for the testbed experiments are the total number n_{total} of peers, the number n_{keys} of peers holding keyshares, the number n_{thres} of partially signed certificates required for interpolation of a valid certificate, and the probability p_{rep} with which a single peer answers a request in an acceptable time. We combined the factors for the individual experiments such that we obtain reasonable variations of n_{keys} high/low subject to n_{total}, n_{thres} high/medium/low subject to n_{keys} and p_{rep} high/low subject to both n_{keys} and n_{thres}. The resulting combinations are listed in Table 7.3. For each combination, we performed 40 request/reply interactions.

Note that the number n_{thres} of partially signed certificates required to compute a valid certificate given in the table is used to instantiate the model. This way, we determine the number n_{req} of peers to which a request has to be sent to achieve the targeted success probability $p_{succ} = 0.95$. In addition, to validate the model not only for the targeted success probability $p_{succ} = 0.95$, but for any success probability $0 \leq p_{succ} \leq 1$, we varied the threshold in $2 \leq n_{thres} \leq n_{keys}$ for the results presented in the following. That is, after having performed the experiments, we determine how many of the 40 request/reply interactions would have been successful if we would have used a different number of partially signed certificates required to compute a valid certificate.

Table 7.3: Combinations of the factors as used in the testbed experiments

n_{total}	n_{keys}	n_{thres}	p_{rep}
30	15	5	0.5
100	15	10	0.95
100	50	10	0.5
100	80	60	0.95

7.3.9 Analysis of the Results

In the following, we present selected results of the testbed studies that show all effects we observed. All plots are given with 95% confidence intervals.

For each interaction scheme, we first discuss general observations followed by a discussion of the individual experiments. For each experiment, we first compare the predictions of a naïve model instantiation with the experimental results. That is, we first assume an ideal peer-to-peer system without effects such as packet loss and churn. Afterwards, we adapt the model instantiation to the particular network parameters we measured during the experiment. We always present results for both the G-Lab and the PlanetLab testbed. For selected scenarios, we present results for experiments conducted in both testbeds simultaneously. For this, we established a peer-to-peer network consisting of nodes taken from G-Lab as well as PlanetLab.

Where applicable, we include the lower bound approximation of the models derived in Section 7.2.3.

We now compare model predictions and testbed results for the interaction scheme for known shareholders. To analyze the general validity of the model as far as possible, we compare the model predictions with the implementation based on a coordinating peer and based on Scribe as discussed in Section 7.1.1.

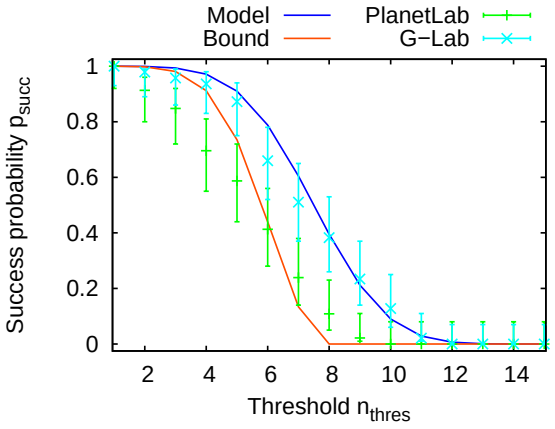
Coordinated Approach

The results obtained for the implementation of the interaction scheme for known shareholders based on a coordinating peer are shown in Figures 7.8 and 7.9.

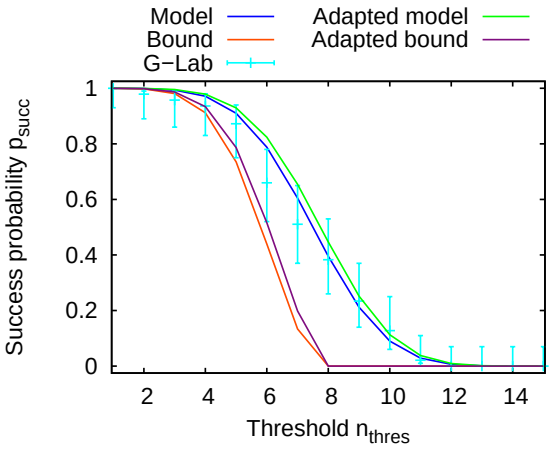
For both setups, we observe that the success probability p_{succ} achieved in PlanetLab is significantly below the success probability observed in G-Lab. The reason for this is the high load of PlanetLab and the resulting churn and inherent packet loss and delay. This causes lost request messages as shown in the tables listing the model adaptations. It stands out that, on the other hand, the probability p_{rep} with which a single peer provides a reply to a request received is not affected significantly. We conclude that the lost request messages are mainly caused by the peer that acts as mediator for the interaction of requesting peer and peers holding keyshares. If this single point of failure is unavailable, request messages are not relayed properly and the decision process cannot be completed.

Regarding the accuracy of the model, we see that the predictions of the naïve model instantiation match the results obtained from G-Lab reasonably. As can be seen in the tables, the network conditions measured in the testbed only show minor deviations from the naïve instantiation. Thus, we observe only a minor deviation of naïve model predictions and predictions of the model adapted to the conditions of G-Lab. For PlanetLab, we observe stronger deviations, in particular for the scenario with a total number of $n_{total} = 100$ peers shown in Figure 7.9. Here, the coordinating peer is affected heavily resulting in that, in average, only 49 of the 65 requests required to achieve a success probability of $p_{succ} = 0.95$ can be delivered. This is further amplified by the fact that, due to churn, only 69 of 80 peers holding keyshares in the original scenario design were available in average.

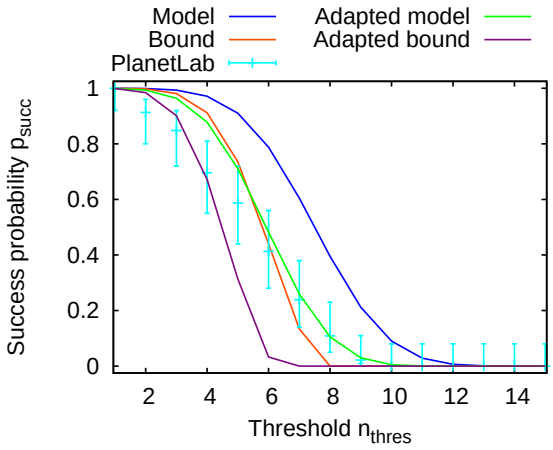
Regarding the closed-form description of the models, we see that, for G-Lab, already the naïve instantiation provides a correct lower bound for the success probability. For PlanetLab, the closed-form provides correct results in the scenario with $n_{total} = 30$ peers, shown in Figure 7.8, although the probability $p_{rep} = 0.49$ with which a single peer provides a reply is slightly below the range in which the Chernoff bound is applicable.



(a) Comparison of naïve model predictions and experimental results from the PlanetLab and G-Lab testbeds



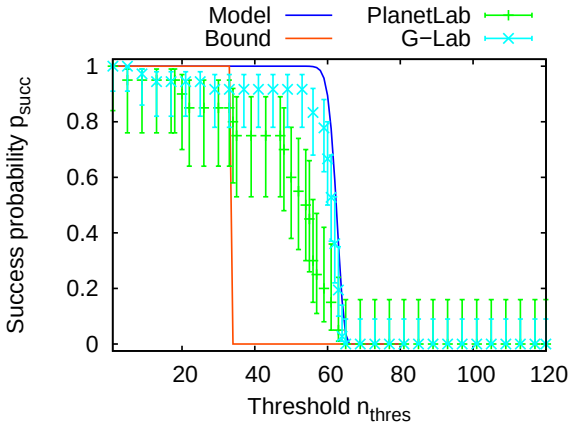
(b) Comparison of naïve model predictions and adapted model for the G-Lab testbed



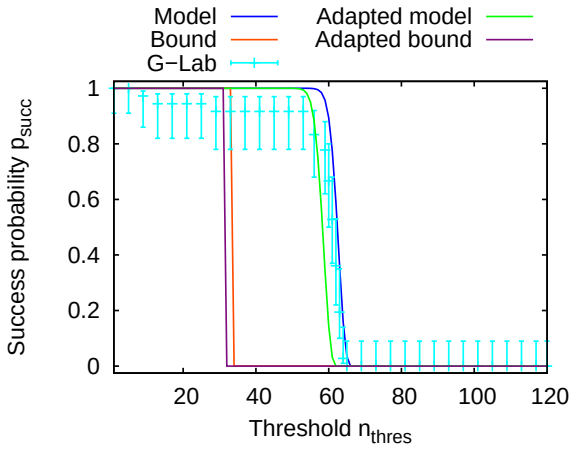
(c) Comparison of naïve model predictions and adapted model for the PlanetLab testbed

	Naïve model	G-Lab adaptation	PlanetLab adaptation
n_{total}	30	30	28
n_{keys}	15	15	12
p_{rep}	0.5	0.52	0.49
n_{req}	14	14	11

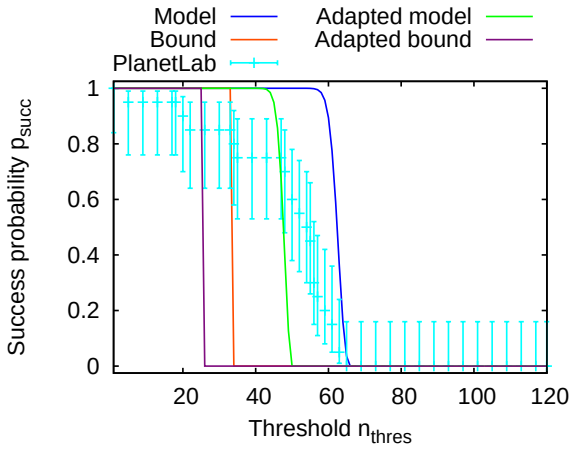
Figure 7.8: Success probability subject to threshold for the coordinated approach. Comparison of model predictions and experimental results for a setup with 30 peers, 15 keyshares and threshold 5. Corresponding naïve model instantiations and instantiations adapted to the conditions observed in the testbeds are given in the table.



(a) Comparison of naïve model predictions and experimental results from the PlanetLab and G-Lab testbeds



(b) Comparison of naïve model predictions and adapted model for the G-Lab testbed



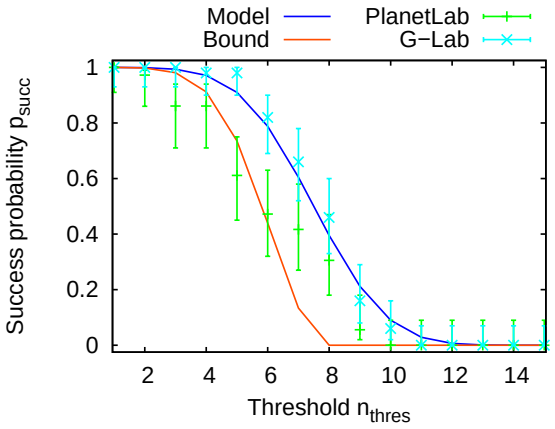
(c) Comparison of naïve model predictions and adapted model for the PlanetLab testbed

	Naïve model	G-Lab adaptation	PlanetLab adaptation
n_{total}	100	100	92
n_{keys}	80	79	68
p_{rep}	0.95	0.94	0.96
n_{req}	65	61	49

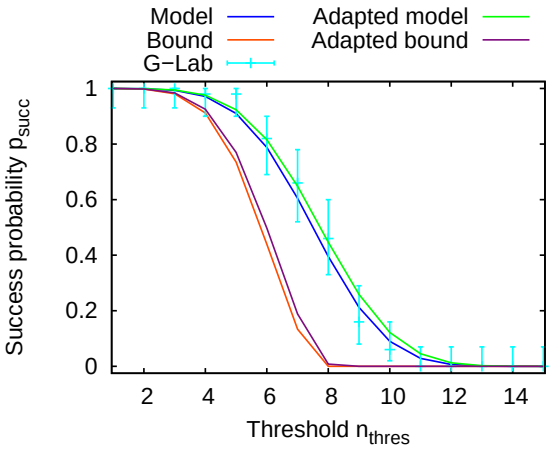
Figure 7.9: Success probability subject to threshold for the coordinated approach. Comparison of model predictions and experimental results for a setup with 100 peers, 80 keyshares and threshold 60. Corresponding naïve model instantiations and instantiations adapted to the conditions observed in the testbeds are given in the table.

Scribe

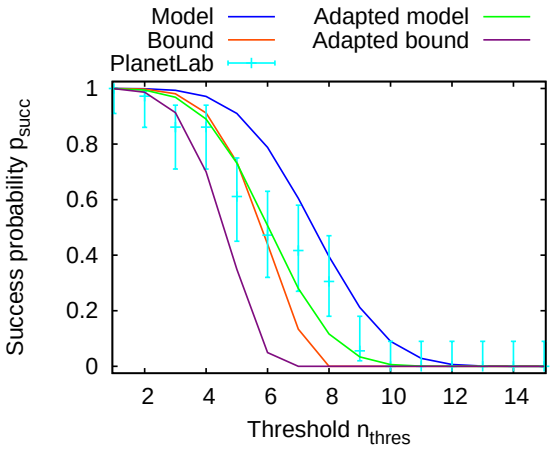
Figures 7.10 to 7.12 show the results for the implementation of the interaction scheme for known shareholders based on Scribe multicast groups.



(a) Comparison of naïve model predictions and experimental results from the PlanetLab and G-Lab testbeds



(b) Comparison of naïve model predictions and adapted model for the G-Lab testbed

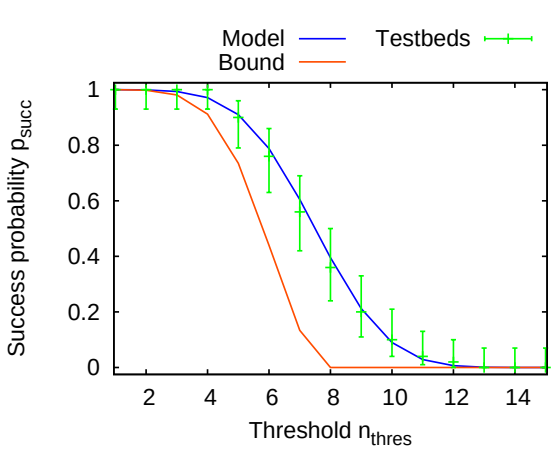


(c) Comparison of naïve model predictions and adapted model for the PlanetLab testbed

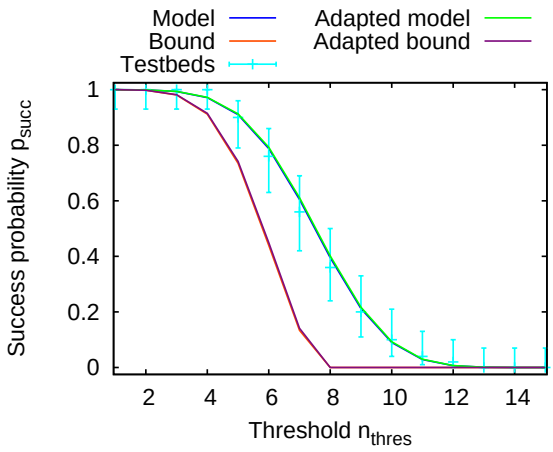
	Naïve model	G-Lab adaptation	PlanetLab adaptation
n_{total}	30	31	29
n_{keys}	15	15	12
p_{rep}	0.5	0.48	0.5
n_{req}	14	15	11

Figure 7.10: Success probability subject to threshold for the multicast approach based on Scribe. Comparison of model predictions and experimental results for a setup with 30 peers, 15 keyshares and threshold 5. Corresponding naïve model instantiations and instantiations adapted to the conditions observed in the testbeds are given in the table.

As for the approach based on a mediating peer described in the previous section, we observe that the success probability is significantly lower for experiments conducted in PlanetLab than for experiments



(a) Comparison of naïve model predictions and experimental results for nodes taken from PlanetLab and G-Lab simultaneously



(b) Comparison of naïve model predictions and adapted model for nodes taken from PlanetLab and G-Lab simultaneously

	Naïve model	Testbed adaptation
n_{total}	30	30
n_{keys}	15	14
p_{rep}	0.5	0.5
n_{req}	14	14

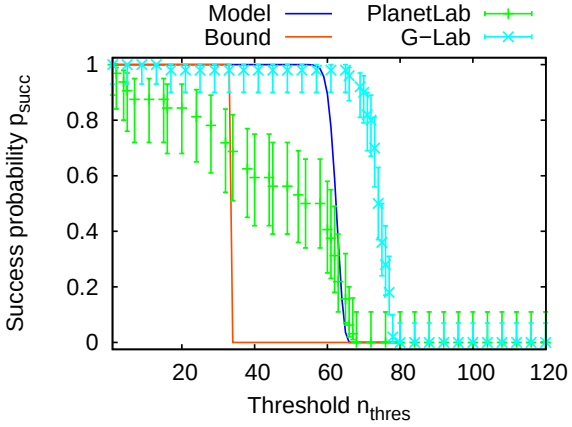
Figure 7.11: Success probability subject to threshold for the multicast approach based on Scribe. Comparison of model predictions and experimental results for a setup with 30 peers, 15 keyshares, threshold 5, and nodes taken from PlanetLab and G-Lab simultaneously. Corresponding naïve model instantiations and instantiations adapted to the conditions observed in the testbeds are given in the table.

conducted in G-Lab. Again, the probability p_{rep} with which a single peer provides a reply to a request received is not affected significantly. Thus, delivery of packets containing partially signed certificates works properly also in G-Lab. Yet, especially in the scenario with a total number of $n_{total} = 100$ peers, both the number n_{req} of requests delivered correctly and the number n_{keys} of peers holding keyshares that are available in the multicast tree are degraded strongly compared to the scenario design. Thus, the churn observed in PlanetLab affects both the entry point and the stability of the multicast tree. As can be deduced from Figure 7.11, the tree can be stabilized if G-Lab peers are mixed among peers set up on PlanetLab nodes.

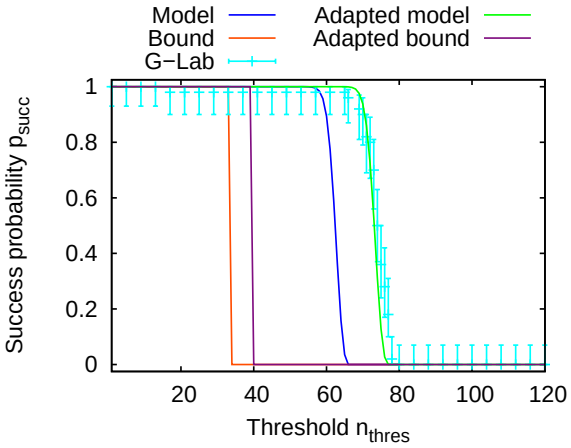
Regarding model accuracy, the naïve instantiation matches the results obtained from G-Lab and from both testbeds simultaneously without significant deviations for the scenario with a total number of $n_{total} = 30$ peers shown in Figures 7.10 and 7.11. It stands out that, for the scenario with $n_{total} = 100$ peers, the results obtained from G-Lab are better than predicted by the model. This is caused by the high stability of the multicast tree and its entry point. 78 of the 80 peers holding keyshares are available in the multicast tree. As for the interaction based on a mediating peer presented in the previous section, we observe that the results obtained from PlanetLab match the adapted model reasonably for the scenario consisting of $n_{total} = 30$ peers. For the scenario consisting of $n_{total} = 100$ peers, the multicast tree cannot be stabilized properly. Thus, a high success probability cannot be achieved. Still, for a low success probability of $p_{succ} < 0.4$, model predictions and testbed results match.

Regarding the closed-form description, the lower bound holds already in the naïve instantiation for experiments conducted in G-Lab and with nodes taken from both G-Lab and PlanetLab. Also for the

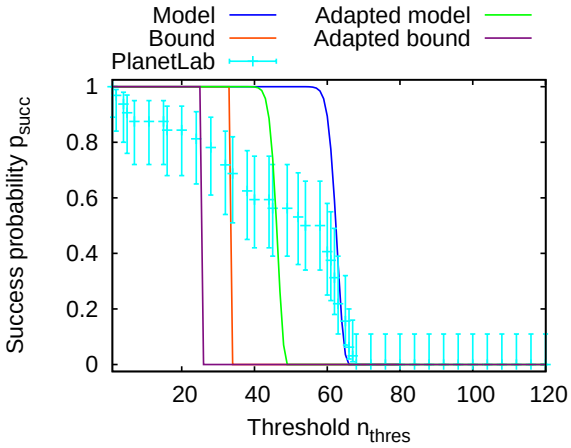
scenario consisting of $n_{total} = 30$ peers taken from PlanetLab, the Chernoff bound is a reasonable estimate of the success probability already for the naïve instantiation. Due to the unstable multicast tree, we cannot rate the quality of the bound for the scenario consisting of $n_{total} = 100$ peers taken from PlanetLab.



(a) Comparison of naïve model predictions and experimental results from the PlanetLab and G-Lab testbeds



(b) Comparison of naïve model predictions and adapted model for the G-Lab testbed



(c) Comparison of naïve model predictions and adapted model for the PlanetLab testbed

	Naïve model	G-Lab adaptation	PlanetLab adaptation
n_{total}	100	100	95
n_{keys}	80	78	51
p_{rep}	0.95	0.95	0.95
n_{req}	65	76	48

Figure 7.12: Success probability subject to threshold for the multicast approach based on Scribe. Comparison of model predictions and experimental results for a setup with 100 peers, 80 keyshares and threshold 60. Corresponding naïve model instantiations and instantiations adapted to the conditions observed in the testbeds are given in the table.

We now present selected results for the interaction scheme for unknown shareholders. As for the interaction scheme for known shareholders, we present results obtained for two different implementations based on requests sent to randomly selected peer IDs and on a gossiping of requests in the Pastry leafset as discussed in Section 7.1.2.

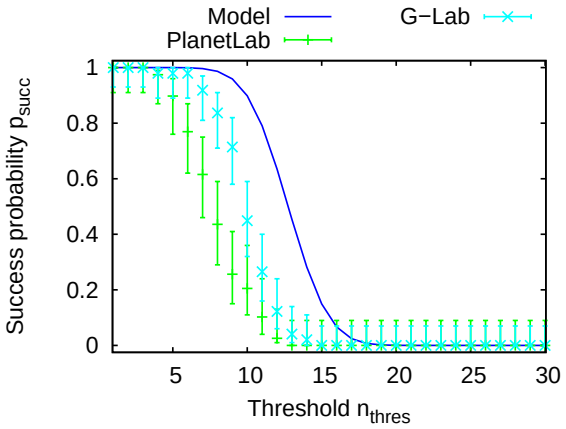
Random Shooting

The results obtained for the interaction of requesting peer and peers that contribute to a decision process based on randomly selected peer IDs are shown in Figures 7.13 to 7.15.

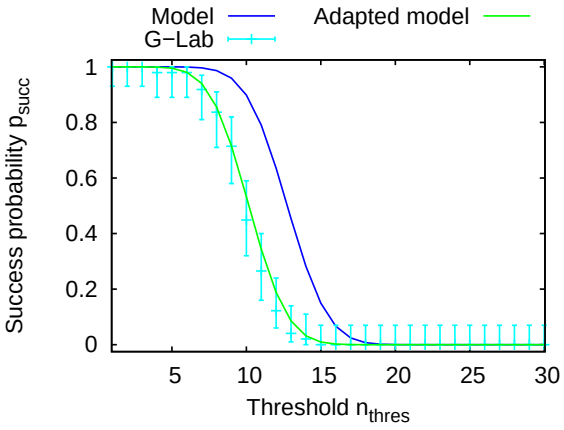
As observed for the interaction schemes for known shareholders discussed in the previous section, the success probability p_{succ} measured in PlanetLab is significantly lower than the success probability measured in G-Lab. This, again, is due to the inherent heterogeneity and the high load of the testbed which results in an increased number of partially unavailable peers. In contrast to the network conditions measured in the testbeds so far, it stands out that, for the random shooting interaction scheme, the probability p_{rep} with which a single peer provides a reply to a request received is degraded significantly compared to the experimental design. We put this down to duplicate requests. If a request is shot randomly to a peer that does not hold a keyshare or to a non-existing peer ID, the request is routed to the peer holding a keyshare with the peer ID numerically closest to the one the request was initially sent to. Thus, requests that are sent to different randomly selected peer IDs may reach the same peer holding a keyshare. Since multiple partial signatures that are produced with the same keyshare are of no use for interpolation of a valid certificate, we drop these duplicate requests. In this case, the probability p_{rep} with which the peer that received multiple requests provides a reply to a request received degrades.

Regarding model accuracy, churn and duplicate requests lead to an overestimate of the success probability p_{succ} for the naïve model instantiation with respect to both G-Lab and PlanetLab. Yet, in contrast to the results obtained for the interaction schemes with known shareholders, we observe a reasonable match of the models adapted to the conditions in G-Lab, PlanetLab, and for a peer-to-peer system consisting of nodes taken from both testbeds. We conclude that, due to not depending on dedicated peers, the random shooting interaction scheme is more robust to packet loss and churn.

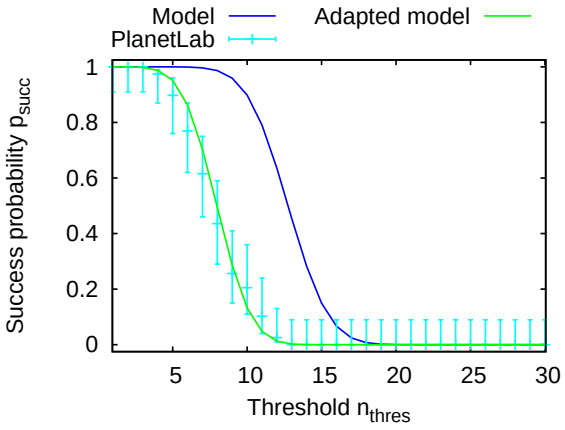
Regarding the closed-form model, the Chernoff bound is not applicable in the scenario with a total number of $n_{total} = 100$ peers and $n_{keys} = 50$ shareholders shown in Figure 7.13. This is due to the probability p_{rep} with which a single peer provides a reply to a request received that, caused by duplicate requests, is significantly below 0.5 in this scenario. Also for the scenarios consisting of $n_{total} = 100$ peers of which $n_{keys} = 80$ hold keyshares, shown in Figures 7.14 and 7.15, we observe that the naïve instantiation of the Chernoff bound overestimates the success probability p_{succ}. This is caused by the fact that, in these scenarios, a high number n_{req} of requests with respect to the total number of peers n_{total} is sent. Thus, we get a high error of the binomial approximation for the hypergeometric random variable which leads to the incorrect naïve bound. Still, the instantiations of the bound that are adapted to the network conditions observed provide correct results.



(a) Comparison of naïve model predictions and experimental results from the PlanetLab and G-Lab testbeds



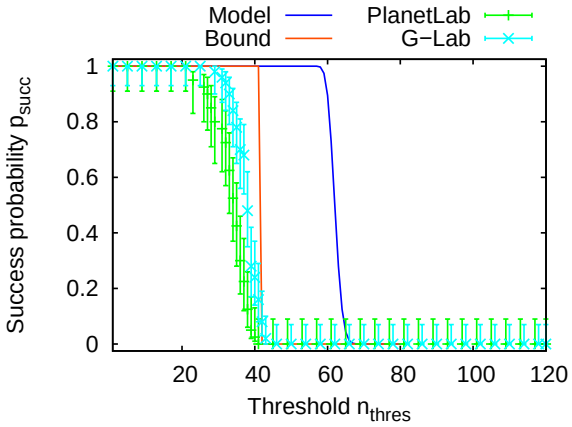
(b) Comparison of naïve model predictions and adapted model for the G-Lab testbed



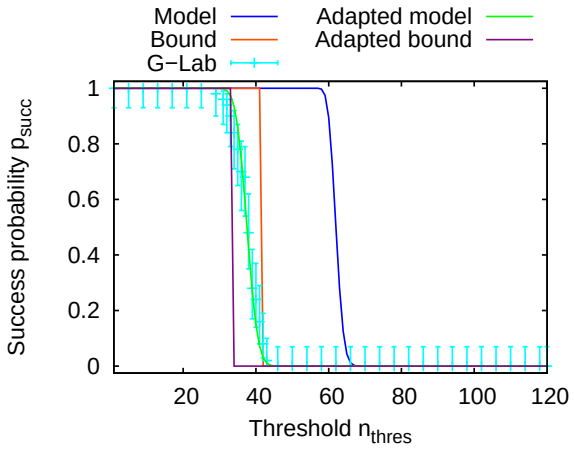
(c) Comparison of naïve model predictions and adapted model for the PlanetLab testbed

	Naïve model	G-Lab adaptation	PlanetLab adaptation
n_{total}	100	100	92
n_{keys}	50	50	39
p_{rep}	0.5	0.43	0.4
n_{req}	49	44	43

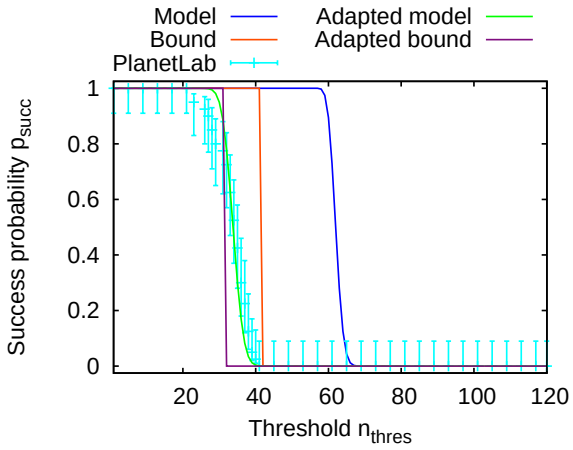
Figure 7.13: Success probability subject to threshold for the random shooting approach. Comparison of model predictions and experimental results for a setup with 100 peers, 50 keyshares and threshold 10. Corresponding naïve model instantiations and instantiations adapted to the conditions observed in the testbeds are given in the table.



(a) Comparison of naïve model predictions and experimental results from the PlanetLab and G-Lab testbeds



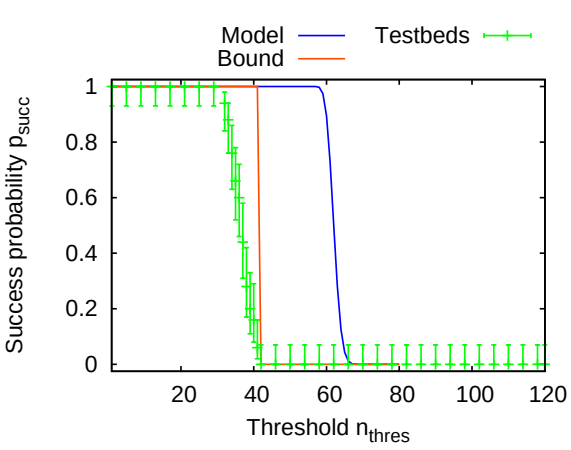
(b) Comparison of naïve model predictions and adapted model for the G-Lab testbed



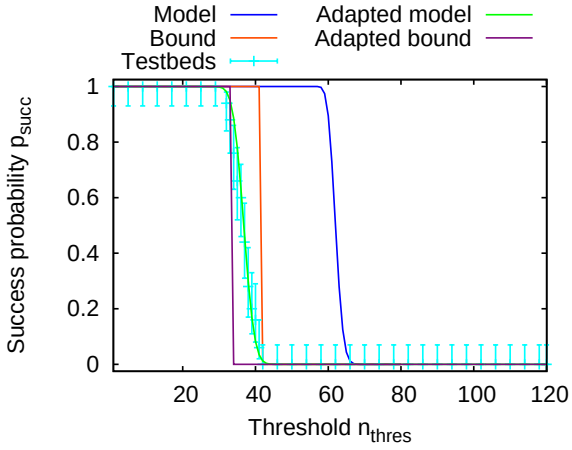
(c) Comparison of naïve model predictions and adapted model for the PlanetLab testbed

	Naïve model	G-Lab adaptation	PlanetLab adaptation
n_{total}	100	100	95
n_{keys}	80	79	68
p_{rep}	0.95	0.72	0.75
n_{req}	81	65	62

Figure 7.14: Success probability subject to threshold for the random shooting approach. Comparison of model predictions and experimental results for a setup with 100 peers, 80 keyshares and threshold 60. Corresponding naïve model instantiations and instantiations adapted to the conditions observed in the testbeds are given in the table.



(a) Comparison of naïve model predictions and experimental results for nodes taken from PlanetLab and G-Lab simultaneously



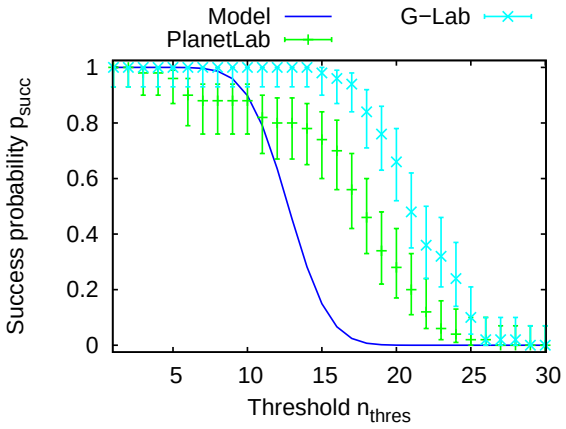
(b) Comparison of naïve model predictions and adapted model for nodes taken from PlanetLab and G-Lab simultaneously

	Naïve model	Testbed adaptation
n_{total}	100	97
n_{keys}	80	74
p_{rep}	0.95	0.74
n_{req}	81	64

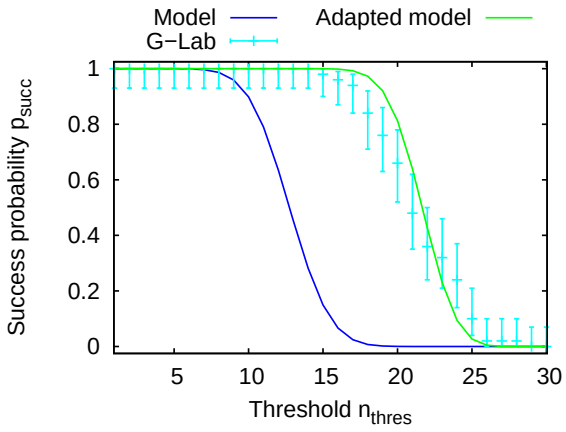
Figure 7.15: Success probability subject to threshold for the random shooting approach. Comparison of model predictions and experimental results for a setup with 100 peers, 80 keyshares, threshold 60, and nodes taken from PlanetLab and G-Lab simultaneously. Corresponding naïve model instantiations and instantiations adapted to the conditions observed in the testbeds are given in the table.

Gossiping

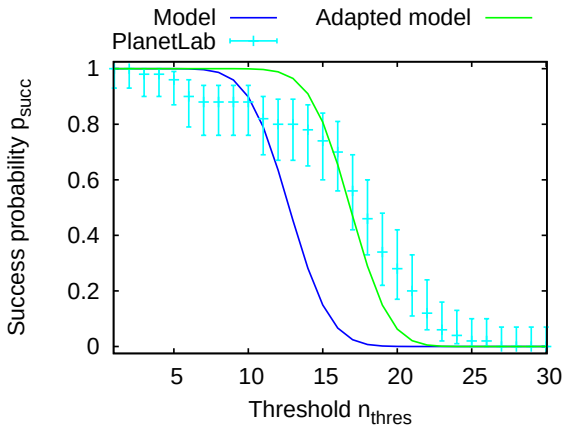
The results for the implementation of the interaction scheme for unknown shareholders based on gossiping a request to logical neighbors are shown in Figures 7.16 to 7.18.



(a) Comparison of naïve model predictions and experimental results from the PlanetLab and G-Lab testbeds



(b) Comparison of naïve model predictions and adapted model for the G-Lab testbed



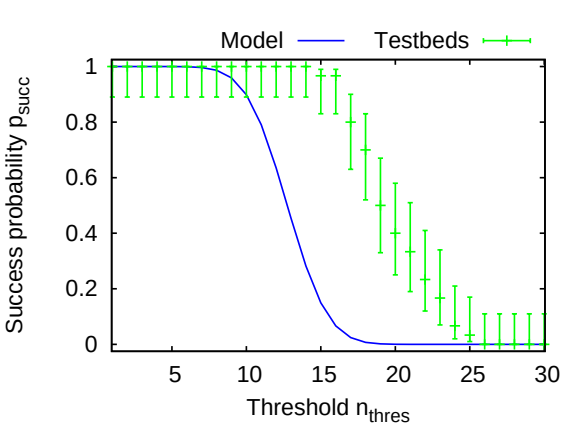
(c) Comparison of naïve model predictions and adapted model for the PlanetLab testbed

	Naïve model	G-Lab adaptation	PlanetLab adaptation
n_{total}	100	101	94
n_{keys}	50	50	38
p_{rep}	0.5	0.53	0.68
n_{req}	49	79	59

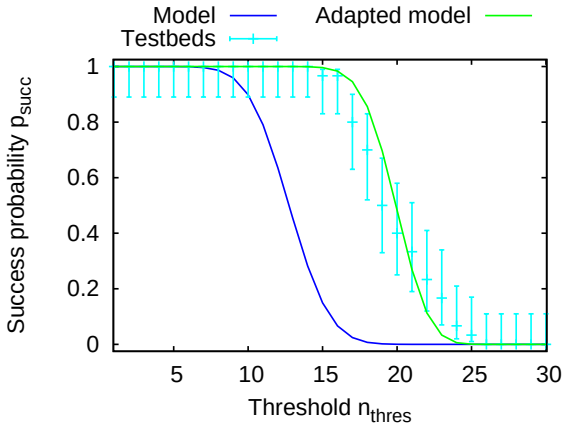
Figure 7.16: Success probability subject to threshold for the gossiping approach. Comparison of model predictions and experimental results for a setup with 100 peers, 50 keyshares and threshold 10. Corresponding naïve model instantiations and instantiations adapted to the conditions observed in the testbeds are given in the table.

As for all experiments presented so far, the success probability p_{succ} observed in PlanetLab is significantly below the success probability observed in G-Lab due to temporarily unavailable peers caused by the high load of the testbed. In contrast to the random shooting approach discussed previously, the prob-

ability p_{rep} with which a single peer provides a reply to a request received is not degraded, but increased compared to the scenario design. This is due to the fact that if a peer holding a keyshare receives a duplicate request, in the gossiping implementation, the request is not dropped, but forwarded further. Except for the scenario consisting of $n_{total} = 100$ peers of which $n_{keys} = 80$ hold keyshares, we also observe an increased number of requested peers than required to achieve the targeted success probability of $p_{succ} = 0.95$. This is due to the fact that the counter indicating the remaining number of partially signed certificates required is rounded to the next higher integer when the request is sent to the leafset of the initiating peer. Considering, for example, the scenario consisting of $n_{total} = 100$ peers of which $n_{keys} = 50$ hold keyshares, the model demands to send $n_{req} = 49$ requests to achieve the targeted success probability of $p_{succ} = 0.95$. For a typical Pastry leafset consisting of 24 peers, the counter indicating the number of partial signatures required should be set to $49 \cdot 24^{-1} = 2.04$, but is rounded to 3.



(a) Comparison of naïve model predictions and experimental results for nodes taken from PlanetLab and G-Lab simultaneously



(b) Comparison of naïve model predictions and adapted model for nodes taken from PlanetLab and G-Lab simultaneously

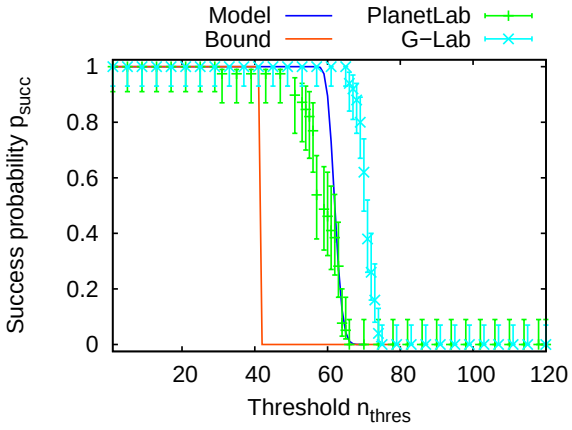
	Naïve model	Testbed adaptation
n_{total}	100	89
n_{keys}	50	45
p_{rep}	0.5	0.55
n_{req}	49	69

Figure 7.17: Success probability subject to threshold for the gossiping approach. Comparison of model predictions and experimental results for a setup with 100 peers, 50 keyshares, threshold 10, and nodes taken from PlanetLab and G-Lab simultaneously. Corresponding naïve model instantiations and instantiations adapted to the conditions observed in the testbeds are given in the table.

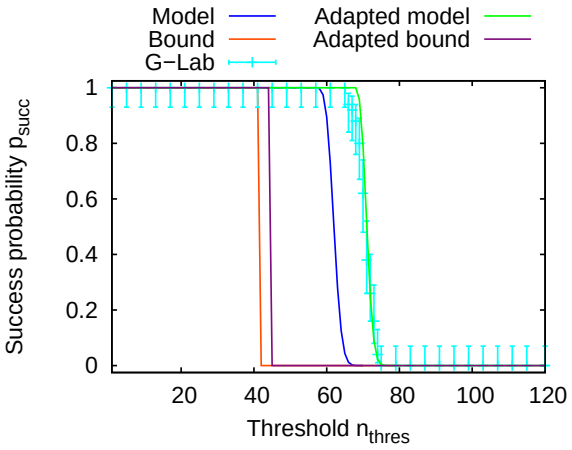
Regarding model accuracy, we observe that the naïve model instantiation underestimates the success probability p_{succ} measured in the testbeds except for the experiments conducted in PlanetLab for the scenario consisting of $n_{total} = 100$ peers of which $n_{keys} = 80$ hold keyshares. In this case, the packet loss and churn of PlanetLab outbalance the increased number of requests initiated due to rounding effects. For all other experiments, the increased success probability compared to the model predictions can be explained by duplicate requests that are forwarded instead of being dropped and an increased number of request messages due to rounding effects.

Regarding the closed-form, the binomial approximation to the hypergeometric random variable and, thus, the Chernoff bound is not applicable for the scenario consisting of $n_{total} = 100$ peers of which

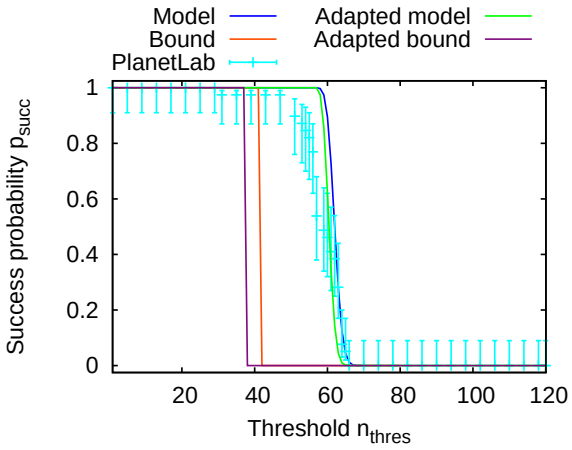
$n_{keys} = 50$ hold keyshares shown in Figures 7.16 and 7.17. For the scenario consisting of $n_{total} = 100$ peers of which $n_{keys} = 80$ hold keyshares shown in Figure 7.18, the closed-form provides a correct lower bound.



(a) Comparison of naïve model predictions and experimental results from the PlanetLab and G-Lab testbeds



(b) Comparison of naïve model predictions and adapted model for the G-Lab testbed



(c) Comparison of naïve model predictions and adapted model for the PlanetLab testbed

	Naïve model	G-Lab adaptation	PlanetLab adaptation
n_{total}	100	101	90
n_{keys}	80	80	64
p_{rep}	0.95	0.99	0.99
n_{req}	81	87	72

Figure 7.18: Success probability subject to threshold for the gossiping approach. Comparison of model predictions and experimental results for a setup with 100 peers, 80 keyshares and threshold 60. Corresponding naïve model instantiations and instantiations adapted to the conditions observed in the testbeds are given in the table.

7.4 Conclusion - User-based Cooperative Decisions

In this chapter, we discussed user-based cooperative decisions in peer-to-peer systems. By means of threshold cryptography, we can compensate missing central trusted instances in spontaneously established peer-to-peer systems. By involving users directly in security-related decisions, we can compensate missing security policies.

When users are involved in the decision process, we cannot assume that a user is available for contributing to a decision at any time. The threshold cryptography approach deployed in our scenario must be able to deal with this. We identified non-interactive variants of threshold cryptography as appropriate solutions. Here, no knowledge about which shareholders contribute to a decision process is required in advance, among the shareholders and at the peer that requests a security-related decision. As soon as a sufficiently large number of partially signed certificates is collected by the requesting peer, a valid certificate can be interpolated.

For the interaction between requesting peer and peers that potentially contribute to a decision process, we discussed different interaction schemes. To be able to control the number of users involved in a decision, we provided mathematical models describing the interaction schemes. Since the exact conditions such as packet loss and churn in a peer-to-peer system might not be known, we additionally derived a lower bound for the success probability (that is, the probability that a sufficiently large number of partially signed certificates can be collected from one request) of an interaction scheme subject to the number of users requested.

In a series of experiments in two different testbeds, we verified the general applicability of the models developed. Although we partially observed deviations of predictions made by a naïve model instantiation assuming an ideal network and the results measured in testbeds, predictions of model instantiations adapted to the conditions of the testbeds match the testbed results reasonably. Where applicable, the lower bound provided correct estimates without requiring an adaptation to the network conditions.

8 Conclusions

In this thesis, we scrutinized how two essential security services, intrusion response and security-related decisions, can be realized in networks operating in absence of a communication infrastructure and without central, trusted instances. We considered the application scenario of a communication network based on mobile ad hoc networks and peer-to-peer applications, as it might be established spontaneously in large-scale disaster scenarios. We assumed that the network is established scenario-wide, consisting of any devices, of affected civilians as well as of first responders and reconstruction units, that are equipped with wireless communication capabilities. Within such a network, mechanisms to deal with malfunctioning devices or intended misbehavior are required. We focused on the question of how to exclude such devices from the network efficiently, once detected. We further assumed that predefined closed user groups exist within the network, but that dynamic inter-group access to restricted resources should be possible. The second focus of the thesis was set on how corresponding security objectives like authentication and access control can be realized without central, trusted instances and security policies.

To exclude misbehaving devices from the network, we developed an intrusion response mechanism that operates based on quarantined areas that are established at locations where misbehavior is detected. This way, we render the intrusion response mechanism insusceptible to changes in addresses of misbehaving nodes. From our point of view, this is a basic step, since preventing to be identified and penalized by changing addresses is possible with little effort but causes a high impact in a scenario where devices are beyond a central control. Being independent from addresses further enables reacting to adverse situations in which the origin of a network interference cannot be traced-back to a particular network address. A sophisticated jamming of radio frequencies, for example, may considerably degrade the performance of routes passing through the jammed area, thus affecting also devices that are not located in radio range of the jamming device.

We evaluated the location-based intrusion response in a series of simulation studies. Our objective was to compare the performance to an address-based solution. To show the effects of changes in addresses of misbehaving nodes, we combined a black hole attack with a Sybil attack. Our core findings show that, in this scenario, the location-based intrusion response is clearly superior. It can be used to support preventive security mechanisms based on means of cryptography. However, it is afflicted with inherent drawbacks. These result from quarantined benign nodes that are excluded from the network by the location-based intrusion response because of their proximity to misbehaving nodes.

In order to be able to predict system behavior for scenarios different from the ones we studied by means of simulation, we developed an analytical model that describes the performance of the location-based intrusion response. We mutually validated model predictions and results obtained from simulation studies.

To improve the location-based intrusion response, we designed and evaluated two approaches. First, we investigated to what extent an adaptive transmission power can be used to reduce the size of quarantined areas. Our results obtained in simulation studies showed that the negative effects of the location-based intrusion response caused by quarantined benign nodes can be diminished effectively. Yet, the routing protocol we used for the evaluation was not designed for an adaptive transmission power. In particular in scenarios with a high node mobility, we observed side-effects that caused frequent route breaks which rendered an adaptive transmission power infeasible in this cases.

As second approach for improving the location-based intrusion response, we enabled delayed communication by harnessing the mobility of quarantined benign nodes. For this, data that has to be sent by a quarantined node is buffered until the node left quarantine. Results based on simulation studies showed that this approach can nearly fully recover network performance in terms of the packet delivery ratio, even in presence of multiple misbehaving nodes. On the downside, the transmission delay measured

from source to destination is increased significantly. Thus, a delayed communication can be used for applications with weak time constraints.

For the second focus of this thesis, that is, for achieving security objectives like authentication and access control without central, trusted instances and security policies, we designed means for user-based cooperative decisions. By involving authorized users directly, we enable security related decisions without requiring predefined security policies. By means of threshold cryptography, we can enforce a cooperative decision process. This way, we prevent that single, unauthorized users holding compromised devices can decide on security-related requests unrestrictedly.

When involving users in a decision process, we have to take into account users that are not able to contribute to a cooperative decision in a reasonable time. Thus, a security-related request has to be sent to a sufficiently large number of users, taking into account potential missing replies. On the other hand, the number of users requested per decision should be minimized. Considering that the main tasks of users in disaster scenarios are, most likely, different ones, involving a user in to many decision processes may degrade the quality of experience and, thus, the feasibility of user-based cooperative decisions.

For controlling the number of users involved in a decision process, we developed analytical models based on different schemes for the interaction between a user that issues a security-related request and users that potentially contribute to the corresponding decision process. We validated the models by controlling a prototype that enables user-based cooperative decisions deployed in two different testbeds. The results proved the applicability of the models for varying network conditions observed in the testbeds.

Well-controllable communication infrastructures and central, trusted instances are key building blocks of security mechanisms for contemporary communication networks. Altogether, the results we presented in this thesis show that, also without these key building blocks, a reasonable level of security for infrastructure-less and decentralized communication networks can be achieved.

Bibliography

[1] Imad Aad, Jean-Pierre Hubaux, and Edward W. Knightly. Impact of Denial of Service Attacks on Ad Hoc Networks. *IEEE/ACM Transactions on Networking (TON)*, 16:791–802, 2008.

[2] Gergely Acs, Levente Buttyan, and Istvan Vajda. Provable Security of On-Demand Distance Vector Routing in Wireless Ad Hoc Networks. In *Proceedings of the 2nd European Workshop on Security and Provacy in Ad hoc and Sensor Networks (ESAS '05)*, 2005.

[3] Gergely Acs, Levente Buttyan, and Istvan Vajda. Provably Secure On-Demand Source Routing in Mobile Ad Hoc Networks. *IEEE Transactions on Mobile Computing*, 5:1533–1546, 2006.

[4] Frooq Anjum, Dhanant Subhadrabandhu, and Saswati Sarkar. Signature based Intrusion Detection for Wireless Ad-Hoc Networks: A Comparative study of various routing protocols. In *Proceedings of the 58th IEEE Vehicular Technology Conference (VTC '03 Fall)*, 2003.

[5] Fan Bai, Narayanan Sadagopan, Bhaskar Krishnamachari, and Ahmed Helmy. Modeling Path Duration Distributions in MANETs and their Impact on Reactive Routing Protocols. *IEEE Journal on Selected Areas in Communications*, 22:1357–1373, 2004.

[6] Hari Balakrishnan, Srinivasan Seshan, Elan Amir, and Randy H. Katz. Improving TCP/IP Performance over Wireless Networks. In *Proceedings of the 1st Annual International Conference on Mobile Computing and Networking (MobiCom '95)*, 1995.

[7] Sorav Bansal and Mary Baker. Observation-based Cooperation Enforcement in Ad hoc Networks. Technical report, Stanford University, 2003.

[8] Rimon Barr. *An Efficient, Unifying Approach to Simulation using Virtual Machines*. PhD thesis, Cornell University, 2004.

[9] Stefano Basagni, Imrich Chlamtac, Violet R. Syrotiuk, and Barry A. Woodward. A Distance Routing Effect Algorithm for Mobility (DREAM). In *Proceedings of the 4th Annual ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom '98)*, 1998. ISBN 1-58113-035-X.

[10] Douglas M. Blough, Mauro Leoncini, Giovanni Resta, and Paolo Santi. The K-Neigh Protocol for Symmetric Topology Control in Ad Hoc Networks. In *Proceedings of the 4th ACM International Symposium on Mobile Ad Hoc Networking & Computing (MobiHoc '03)*, 2003.

[11] Alexandra Boldyreva. Efficient threshold signature, multisignature and blind signature schemes based on the Gap-Diffie-Hellman-group signature scheme. In *Proceedings of the 6th International Workshop on Practice and Theory in Public Key Cryptography (PKC '03)*, 2003.

[12] John Border, Markku Kojo, James H. Griner, Gabriel Montenegro, and Zach D. Shelby. Performance Enhancing Proxies Intended to Mitigate Link-Related Degradations. *IETF RFC 3135*, 2001.

[13] Sonja Buchegger. *Coping with Misbehavior in Mobile Ad-hoc Networks*. PhD thesis, École Polytechnique Fédérale de Lausanne, 2004.

[14] John Burgess, George Dean Bissias, Mark D. Corner, and Brian Neil Levine. Surviving Attacks on Disruption-Tolerant Networks without Authentication. In *Proceedings of the 8th ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc '07)*, 2007.

[15] Levente Buttyan and Jean-Pierre Hubaux. *Security and Cooperation in Wireless Networks - Thwarting Malicious and Selfish Behavior in the Age of Ubiquitous Computing*. Cambridge University Press, 2008.

[16] Miguel Castro, Peter Druschel, Anne-Marie Kermarrec, and Antony Rowstron. Scribe: A large-scale and decentralized application-level multicast infrastructure. *IEEE Journal on Selected Areas in Communications*, 20:100–110, 2002.

[17] Liang Dai, Yi Cui, and Yuan Xue. On Scalability of Proximity-Aware Peer-to-Peer Streaming. In *Proceedings of the 26th IEEE International Conference on Computer Communications (INFOCOM '07)*, 2007.

[18] Franca Delmastro. From Pastry to CrossROAD: Cross-layer Ring Overlay for Ad Hoc Networks. In *Proceedings of the Third IEEE International Conference on Pervasive Computing and Communications Workshops (PERCOMW '05)*, 2005.

[19] G-Lab Consortium. G-Lab Project Homepage. *http://www.german-lab.de*, 2010.

[20] Michael Gastpar and Martin Vetterli. On The Capacity Of Wireless Networks: The Relay Case. In *Proceedings of the 21st Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM '02)*, 2002.

[21] Rosario Gennaro, Shai Halevi, Hugo Krawczyk, and Tal Rabin. Threshold RSA for Dynamic and Ad-Hoc Groups. In *Proceedings of the International Conference on the Theory and Application of Cryptographic Techniques (EUROCRYPT '08)*, 2008.

[22] Ivan A. Getting. The Global Positioning System. *IEEE Spectrum*, 30:36–38, 43–47, 1993.

[23] Tom Goff, James Moronski, Dhananjay S. Phatak, and Vipul Gupta. Freeze-TCP: A true end-to-end TCP enhancement mechanism for mobile environments. In *Proceedings of the Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM '00)*, 2000.

[24] Javier Gomez and Andrew T. Campbell. A Case for Variable-Range Transmission Power Control in Wireless Multihop Networks. In *Proceedings of the Twenty-third Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM '04)*, 2004.

[25] Javier Gomez, Andrew T. Campbell, Mahmoud Naghshineh, and Chatschik Bisdikian. PARO: Supporting Dynamic Power Controlled Routing in Wireless Ad Hoc Networks. *Wireless Networks*, 9:443–460, 2003.

[26] Li Gong. Lower Bounds on Messages and Rounds for Network Authentication Protocols. In *Proceedings of the 1st ACM Conference on Computer and Communications Security (CCS '93)*, 1993.

[27] Kalman Graffi, Parag S. Mogre, Matthias Hollick, and Ralf Steinmetz. Detection of Colluding Misbehaving Nodes in Mobile Ad Hoc and Mesh Networks. In *Proceedings of the IEEE Global Communications Conference (GLOBECOM '07)*, 2007.

[28] Robin Groenevelt, Philippe Nain, and Ger Koole. Message Delay in MANET. In *Proceedings of the International Conference on Measurements and Modeling of Computer Systems (SIGMETRICS '05)*, 2005.

[29] Matthias Grossglauser and David Tse. Mobility Increases the Capacity of Ad-hoc Wireless Networks. *IEEE/ACM Transactions on Networking (TON)*, 10:477–486, 2002.

[30] Piyush Gupta and P. R. Kumar. The Capacity of Wireless Networks. *IEEE Transactions on Information Theory*, 46:388–404, 2000.

[31] Zygmunt J. Haas, Jing Deng, Ben Liang, Panagiotis Papadimitratos, and Sajama Sajama. Wireless Ad Hoc Networks. In *Encyclopedia of Telecommunications*. John Wiley, 2002.

[32] Ahmed Hasswa, Mohammad Zulkernine, and Hossam Hassanein. Routeguard: An Intrusion Detection and Response System for Mobile Ad Hoc Networks. In *Proceedings of the IEEE International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob '05)*, 2005.

[33] Wenbo He, Ying Huang, Klara Nahrstedt, and Whay C. Lee. Dandelion: Mobility-assisted Reliable Message Propagation Protocol in MANETs. Technical report, Multimedia Operating System and Networking Group, Department of Computer Science, University of Illinois at Urbana-Champaign (UIUC), 2007.

[34] Wenbo He, Ying Huang, Ravishankar Sathyam, Klara Nahrstedt, and Whay C. Lee. SMOCK: A Scalable Method of Cryptographic Key Management for Mission-critical Wireless Ad Hoc Networks. *IEEE Transactions on Information Forensics and Security (TIFS)*, 4:140–150, 2009.

[35] Baik Hoh and Marco Gruteser. Computer Ecology: Responding to Mobile Worms with Location-Based Quarantine Boundaries. In *Proceedings of the International Workshop on Research Challenges in Security and Privacy for Mobile and Wireless Networks (WSPWN '06)*, 2006.

[36] Matthias Hollick. *Dependable Routing for Cellular and Ad hoc Networks*. PhD thesis, Multimedia Communications Lab (KOM), Technische Universität Darmstadt, 2004.

[37] Matthias Hollick, Jens Schmitt, Christian Seipl, and Ralf Steinmetz. On the Effect of Node Misbehavior in Ad Hoc Networks. In *Proceedings of the IEEE International Conference on Communications (ICC '04)*, 2004.

[38] Yih Chun Hu and Adrian Perrig. A Survey of Secure Wireless Ad Hoc Routing. *IEEE Journal on Security and Privacy*, 2:28–39, 2004.

[39] Yih Chun Hu, Adrian Perrig, and David B. Johnson. Ariadne: A Secure On-Demand Routing Protocol for Ad Hoc Networks. In *Proceedings of the 8th Annual International Conference on Mobile Computing and Networking (MobiCom '02)*, 2002.

[40] Yih Chun Hu, Adrian Perrig, and David B. Johnson. Rushing attacks and defense in wireless ad hoc network routing protocols. In *Proceedings of the 2nd ACM Workshop on Wireless Security (WiSe '03)*, 2003.

[41] Yi An Huang and Wenke Lee. A Cooperative Intrusion Detection System for Ad Hoc Networks. In *Proceedings of the 1st ACM Workshop on Security of Ad Hoc and Sensor Networks (SASN '03)*, 2003.

[42] Ying Huang, Wenbo He, Klara Nahrstedt, and Whay C. Lee. Incident Scene Mobility Analysis. In *Proceedings of the 2008 IEEE International Conference on Technologies for Homeland Security (HST '08)*, 2008.

[43] Ying Huang, Wenbo He, and Klara Nahrstedt. ChainFarm: A Novel Authentication Protocol for High-rate Any Source Probabilistic Broadcast. In *Proceedings of the 6th IEEE International Conference on Mobile Ad Hoc and Sensor Systems (MASS '09)*, 2009.

[44] IABG. HiMoNN - Highly Mobile Netwok Node. *http://www.iabg.de/infokom/fachthemen/himonn_en.php*, 2009.

[45] Raj Jain. *The Art of Computer Systems Performance Analysis: Techniques for Experimental Design, Measurement, Simulation, and Modeling*. Wiley, 1991.

[46] Juan J. Jaramillo and Rayadurgam Srikant. DARWIN: Distributed and Adaptive Reputation mechanism for WIreless ad-hoc networks. In *Proceedings of the 13th Annual ACM International Conference on Mobile Computing and Networking (MobiCom '07)*, 2007.

[47] Stanislaw Jarecki, Nitesh Saxena, and Jeong H. Yi. An Attack on the Proactive RSA Signature Scheme in the URSA Ad Hoc Network Access Control Protocol. In *Proceedings of the 2nd ACM Workshop on Security of Ad Hoc and Sensor Networks (SASN '04)*, 2004.

[48] David B. Johnson and David A. Maltz. Dynamic Source Routing in Ad Hoc Wireless Networks. In *Mobile Computing*. Kluwer Academic Publishers, 1996.

[49] Eszter Kail, Gábor Németh, and Zoltán Richárd Turányi. The Effect of the Transmission Range on the Capacity of Ideal Ad Hoc Networks. In *Proceedings of the Fourth International Symposium on Wireless Personal Multimedia Communications (WPMC '01)*, 2001.

[50] Kanchana Kanchanasut, Thirapon Wongsaardsakul, Manutsiri Chansutthirangkool, Anis Laouiti, Hajime Tazaki, and Khandakar Rashedul Arefin. DUMBO II: A V-2-I Emergency Network. In *Proceedings of the 4th Asian Conference on Internet Engineering (AINTEC '08)*, 2008.

[51] Krishna Kant. An Analytic Model for peer to peer File Sharing Networks. In *Proceedings of the IEEE International Conference on Communications (ICC '03)*, 2003.

[52] Thomas H. Kean, Lee H. Hamilton, Richard Ben-Vensite, Bob Kerrey, Fred F. Fielding, John F. Lehman, Jamie S. Gorelick, Timothy J. Roemer, Slade Gorton, and James R. Thompson. The 9/11 Commission Report. *http://www.9-11commission.gov/report/911Report.pdf*, 2004.

[53] Wolfgang Kiess and Martin Mauve. A Survey on Real-world Implementations of Mobile Ad-hoc Networks. *Elsevier Journal on Ad Hoc Networks*, 5:324–339, 2007.

[54] Young-Bae Ko and Nitin H. Vaidya. Location-Aided Routing (LAR) in Mobile Ad Hoc Networks. In *Proceedings of the 4th Annual ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom '98)*, 1998.

[55] Supriya Krishnamurthy, Sameh El-Ansary, Erik Aurell, and Seif Haridi. An Analytical Study of a Structured Overlay in the Presence of Dynamic Membership. *IEEE/ACM Transactions on Networking (TON)*, 16:814–825, 2008.

[56] Tronje Krop, Michael Bredel, Matthias Hollick, and Ralf Steinmetz. JiST/MobNet: Combined Simulation, Emulation and Real-World Testbed for Ad hoc Networks. In *Proceedings of the Second ACM International Workshop on Wireless Network Testbeds, Experimental evaluation and Characterization (WiNTECH '07)*, 2007.

[57] Olaf Landsiedel, Stefan Götz, and Klaus Wehrle. Towards Scalable Mobility in Distributed Hash Tables. In *Proceedings of the Sixth IEEE International Conference on Peer-to-Peer Computing (P2P '06)*, 2006.

[58] Frank Lehrieder, György Dán, Tobias Hoßfeld, Simon Oechsner, and Vlad Singeorzan. The Impact of Caching on BitTorrent-like Peer-to-peer Systems. In *Proceedings of the IEEE International Conference on Peer-to-Peer Computing (P2P '10)*, 2010.

[59] Colin Lemmon, Siu M. Lui, and Ickjai Lee. Geographic Forwarding and Routing for Ad-Hoc Wireless Network: A Survey. In *Proceedings of the 2009 Fifth International Joint Conference on INC, IMS and IDC (NMC '09)*, 2009.

[60] Jinyang Li, Charles Blake, Douglas S. J. De Couto, Hu Imm Lee, and Robert Morris. Capacity of Ad Hoc Wireless Networks. In *Proceedings of the 7th ACM International Conference on Mobile Computing and Networking (MobiCom '01)*, 2001.

[61] Yung-Ming Li, Yong Tan, and Yong-Pin Zhou. Analysis of Scale Effects in Peer-to-Peer Networks. *IEEE/ACM Transactions on Networking (TON)*, 16:590–602, 2008.

[62] Haiyun Luo, Petros Zerfos, Jiejun Kong, Songwu Lu, and Lixia Zhang. Self-securing Ad Hoc Wireless Networks. In *Proceedings of the Seventh International Symposium on Computers and Communications (ISCC '02)*, 2002.

[63] Haiyun Luo, Jiejun Kong, Petros Zerfos, Songwu Lu, and Lixia Zhang. URSA: Ubiquitous and Robust Access Control for Mobile Ad Hoc Networks. *IEEE/ACM Transactions on Networking*, 12: 1049–1063, 2004.

[64] Balakrishnan S. Manoj and Alexandra H. Baker. Communication Challenges in Emergency Response. *Communications of the ACM*, 50:51–53, 2007.

[65] Sergio Marti, Thomas J. Giuli, Kevin Lai, and Mary Baker. Mitigating Routing Misbehavior in Mobile Ad Hoc Networks. In *Proceedings of the 6th Annual International Conference on Mobile Computing and Networking (MobiCom '00)*, 2000.

[66] Martin Mauve, Jörg Widmer, and Hannes Hartenstein. A Survey on Position-Based Routing in Mobile Ad Hoc Networks. *IEEE Network Magazine*, 15:30–39, 2001.

[67] Pietro Michiardi. *Cooperation enforcement and network security mechanisms for mobile ad hoc networks*. PhD thesis, Ecole nationale supérieure des télécommunications, 2004.

[68] Pietro Michiardi and Refik Molva. Analysis of Coalition Formation and Cooperation Strategies in Mobile Ad Hoc Networks. *Elsevier Journal on Ad Hoc Networks*, 3:193–219, 2005.

[69] Leonard E. Miller. Distribution of Link Distances in a Wireless Network. *Journal of Research of the National Institute of Standards and Technology*, 106:401–412, 2001.

[70] Parag Mogre, Kalman Graffi, Matthias Hollick, and Ralf Steinmetz. A Security Framework for Wireless Mesh Networks. *Wireless Communications and Mobile Computing Special Issue on Architectures and Protocols for Wireless Mesh, Ad Hoc, and Sensor Networks*, published online, 2010.

[71] Parag S. Mogre, Kalman Graffi, Matthias Hollick, and Ralf Steinmetz. AntSec: Securing Organically Growing Wireless Mesh Networks. Technical report, Multimedia Communications Lab (KOM), TU Darmstadt, March 2007.

[72] Parag S. Mogre, Kalman Graffi, Matthias Hollick, and Ralf Steinmetz. AntSec, WatchAnt and AntRep: Innovative Security Mechanisms for Wireless Mesh Networks. In *Proceedings of the 32nd IEEE Conference on Local Computer Networks (LCN '07)*, 2007.

[73] Maithili Narasimha, Gene Tsudik, and Jeong H. Yi. On the Utility of Distributed Cryptography in P2P and MANETs: The Case of Membership Control. In *Proceedings of the 11th IEEE International Conference on Network Protocols (ICNP '03)*, 2003.

[74] Clifford Neuman, Tom Yu, Sam Hartman, and Ken Raeburn. The Kerberos Network Authentication Service (V5). *IETF RFC 4120*, 2005.

[75] Seung-Jong Park and Raghupathy Sivakumar. Load-Sensitive Transmission Power Control in Wireless Ad-hoc Networks. In *Proceedings of the IEEE Global Telecommunications Conference 2002 (GLOBECOM '02)*, 2002.

[76] Seung-Jong Park and Raghupathy Sivakumar. Quantitative Analysis of Transmission Power Control in Wireless Ad-hoc Networks. In *Proceedings of the International Conference on Parallel Processing Workshops (ICPPW '02)*, 2002.

[77] Charles E. Perkins and Elizabeth M. Royer. Ad hoc On-Demand Distance Vector Routing. In *Proceedings of the 2nd IEEE Workshop on Mobile Computing Systems and Applications*, 1999.

[78] Charles E. Perkins, Elizabeth M. Belding-Royer, and Samir R. Das. Ad hoc On-Demand Distance Vector (AODV) Routing. *IETF RFC 3561*, 2004.

[79] Adrian Perrig, Ran Canetti, and Doug Tygar. Efficient Authentication and Signing of Multicast Streams over Lossy Channels. In *Proceedings of the IEEE Symposium on Security and Privacy (S&P '00)*, 2000.

[80] PlanetLab Consortium. PlanetLab Homepage. *http://www.planet-lab.org*, 2010.

[81] Thadpong Pongthawornkamol, Klara Nahrstedt, and Guijun Wang. HybridCast: A Hybrid Probabilitstic/Deterministic Approach for Adjustable Broadcast Reliability in Mobile Wireless Ad Hoc Networks. In *Proceedings of the IEEE International Conference on Communications (ICC '09)*, 2009.

[82] Himabindu Pucha, Saumitra M. Das, and Y. Charlie Hu. Ekta: An Efficient DHT Substrate for Distributed Applications in Mobile Ad Hoc Networks. In *Proceedings of the Sixth IEEE Workshop on Mobile Computing Systems and Applications (WMCSA '04)*, 2004.

[83] Matija Pužar, Jon Andersson, Thomas Plagemann, and Yves Roudier. SKiMPy: A Simple Key Management Protocol for MANETs in Emergency and Rescue Operations. In *Proceedings of the Second European Workshop on Security and Privacy in Ad-hoc and Sensor Networks (ESAS '05)*, 2005.

[84] Manikantan Ramadas, Scott Burleigh, and Stephen Farrell. Licklider Transmission Protocol - Specification. *IRTF Delay Tolerant Networking Research Group Internet-Draft*, 2008.

[85] Ram Ramanathan and Regina Rosales-Hain. Topology Control of Multihop Wireless Networks using Transmit Power Adjustment. In *Proceedings of the Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM '00)*, 2000.

[86] Ranga S. Ramanujan and Sid Kudige Thanh Nguyen. Techniques for Intrusion-resistant Ad Hoc Routing Algorithms (TIARA). In *Proceedings of the DARPA Information Survivability Conference and Exposition (DISCEX '03)*, 2003.

[87] Sheldon M. Ross. *Probability Models*. Academic Press, 2003.

[88] Antony Rowstron and Peter Druschel. Pastry: Scalable, decentralized object location and routing for large-scale peer-to-peer systems. In *Proceedings of the IFIP/ACM International Conference on Distributed Systems Platforms (Middleware '01)*, 2001.

[89] Prince Samar and Stephen B. Wicker. On the Behavior of Communication Links of a Node in a Multi-Hop Mobile Environment. In *Proceedings of the 5th ACM International Symposium on Mobile Ad Hoc Networking and Computing*, 2004.

[90] Nitesh Saxena, Gene Tsudik, and Jeong H. Yi. Admission Control in Peer-to-Peer: Design and Performance Evaluation. In *Proceedings of the 1st ACM Workshop on Security of Ad Hoc and Sensor Networks (SASN '03)*, 2003.

[91] Nitesh Saxena, Gene Tsudik, and Jeong H. Yi. Efficient Node Admission for Short-lived Mobile Ad Hoc Networks. In *Proceedings of the 13TH IEEE International Conference on Network Protocols (ICNP '05)*, 2005.

[92] Flora R. Schreiber. *Sybil*. Warner Books, 1973.

[93] Matthias Schwamborn, Nils Aschenbruck, and Peter Martini. A Realistic Trace-based Mobility Model for First Responder Scenarios. In *Proceedings of the 13th ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems (MSWiM '10) (accepted for presentation)*, 2010.

[94] Keith Scott and Scott Burleigh. Bundle Protocol Specification. *IETF RFC 5050*, 2007.

[95] Adi Shamir. How to Share a Secret. *Communications of the ACM*, 22:612–613, 1979.

[96] Victor Shoup. Practical Threshold Signatures. In *Proceedings of the International Conference on the Theory and Application of Cryptographic Techniques (EUROCRPYT '00)*, 2000.

[97] Suresh Singh, Mike Woo, and Cauligi S. Raghavendra. Power-Aware Routing in Mobile Ad Hoc Networks. In *Proceedings of the 4th Annual ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom '98)*, 1998.

[98] Iana Siomina and Di Yuan. Maximizing Lifetime of Broadcasting in Ad Hoc Networks by Distributed Transmission Power Adjustment. In *Proceedings of the 8th International Conference on Transparent Optical Networks (ICTON '06)*, 2006.

[99] Avinash Srinivasan and Jie Wu. A Survey on Secure Localization in Wireless Sensor Networks. In *Encyclopedia of Wireless and Mobile Communications*. CRC Press, 2007.

[100] Vivek Srivastava, James Neel, Allen B. MacKenzie, Rekha Menon, Luiz A. DaSilva, James E. Hicks, Jeffrey H. Reed, and Robert P. Gilles. Using Game Theory to Analyze Wireless Ad Hoc Networks. *IEEE Communications Surveys and Tutorials*, 7:46–56, 2005.

[101] Staatskanzlei Sachsen-Anhalt. Pressemitteilung Nr.: 160/03. *http://www.hochwasser.feuerwehr-magdeburg.org/index.htm?HW Sommer02 Abschluss-LSA.htm*, 2003.

[102] Ralf Steinmetz and Klaus Wehrle, editors. *Peer-to-Peer Systems and Applications*. Springer, 2005.

[103] Warren L. Stutzman and Gary A. Thiele. *Antenna Theory and Design*. Wiley, 1998.

[104] Andrew S. Tanenbaum. *Computer Networks*. Pearson Education, 2002.

[105] Douglas Thain, Todd Tannenbaum, and Miron Livny. Distributed Computing in Practice: The Condor Experience. *Concurrency and Computation: Practice and Experience*, 17:323–356, 2005.

[106] George Theodorakopoulos and John S. Baras. Malicious Users in Unstructured Networks. In *Proceedings of the 26th IEEE International Conference on Computer Communications (INFOCOM '07)*, 2007.

[107] Chin-Yang Tseng, Poornima Balasubramanyam, Calvin Ko, Rattapon Limprasittiporn, Jeff Rowe, and Karl Levitt. A Specification-based Intrusion Detection System for AODV. In *Proceedings of the 1st ACM Workshop on Security of Ad Hoc and Sensor Networks (SASN '03)*, 2003.

[108] Damla Turgut, Sajal K. Das, and Mainak Chatterjee. Longevity of Routes in Mobile Ad hoc Networks. In *Proceedings of the 54th IEEE Vehicular Technology Conference (VTC '01 Fall)*, 2001.

[109] Long Vu and Klara Nahrstedt. Adaptive Mobility-assisted Data Dissenmination in Mobile Disaster/Recovery Environments. In *Proceedings of the 2007 Military Communications Conference (MILCOM '07)*, 2007.

[110] Dan S. Wallach. A Survey of Peer-to-Peer Security Issues. In *Proceedings of the International Symposium on Software Security (ISSS '02)*, 2002.

[111] Chih-Chiang Wang and Khaled Harfoush. On the Stability-Scalability Tradeoff of DHT Deployment. In *Proceedings of the 26th IEEE International Conference on Computer Communications (INFOCOM '07)*, 2007.

[112] Qiyan Wang, Himanshu Khurana, Ying Huang, and Klara Nahrstedt. Time-Valid One-Time Signature for Time-Critical Multicast Data Authentication. In *Proceedings of the 28th IEEE International Conference on Computer Communications (INFOCOM '09)*, 2009.

[113] Joby Warrick. Crisis Communications Remain Flawed - Despite Promises to Fix Systems, First Responders Were Still Isolated After Katrina. *The Washington Post*, http://www.washingtonpost.com/wp-dyn/content/article/2005/12/09/AR2005120902039.html, 2005.

[114] Anthony D. Wood, John A. Stankovic, and Sang H. Son. JAM: A Jammed Area-Mapping Service for Sensor Networks. In *Proceedings of the 24th IEEE International Real-Time Systems Symposium (RTSS '03)*, 2003.

[115] Bing Wu, Jianmin Chen, Jie Wu, and Mihaela Cardei. A Survey on Attacks and Countermeasures in Mobile Ad Hoc Networks. In *Wireless/Mobile Network Security*. Springer, 2006.

[116] Jeong H. Yi. Energy-Efficient and Non-interactive Self-certification in MANETs. In *Proceedings of the 8th International Symposium on Stabilization, Safety, and Security of Distributed Systems (SSS '06)*, 2006.

[117] Manel G. Zapata and Nadarajah Asokan. Securing Ad hoc Routing Protocols. In *Proceedings of the 3rd ACM Workshop on Wireless Security (WiSE '02)*, 2002.

[118] Yongguang Zhang, Wenke Lee, and Yi-An Huang. Intrusion Detection Techniques for Mobile Wireless Networks. *ACM/Kluwer Journal on Wireless Networks*, 9:545–556, 2003.

[119] Nianjun Zhou and Alhussein Alhussein A. Abouzeid. Information Theoretic Analysis of Proactive Routing Overhead in Mobile Ad Hoc Networks. *IEEE Transactions on Information Theory*, 55: 4608–4625, October 2009.

[120] Yipeng Zhou, Dah-Ming Chiu, and John C. S. Lui. A Simple Model for Chunk-Scheduling Strategies in P2P Streaming. *IEEE/ACM Transactions on Networking (TON) (accepted for publication)*, 2010.

A Notations of Formulae

Notations are listed for the chapter they were first used.

A.1 Chapter 2

P_S	Power emitted by a sending antenna
P_R	Power received by a receiving antenna
S_{iso}	Power density of an isotropic antenna
$A_{eff,iso}$	Effective area of an isotropic antenna
$P_{R,free,iso}$	Power received assuming a free space model and an isotropic antenna
S_{omni}	Power density of an omnidirectional antenna
$A_{eff,omni}$	Effective area of an omnidirectional antenna
$P_{R,free,omni}$	Power received assuming a free space model and an omnidirectional antenna
$P_{R,tworay}$	Power received assuming a two ray ground reflection model
r_C	Cross-over distance
$rreq_{SD}$	Route request for Node D initiated by Node S
ip_X	Address of Node X
hop_X	Distance to Node X in hops
sn_X	Sequence number of Node X
$rrep_{DS}$	Route reply for $rreq_{SD}$
$rerr_D$	Route error message for Destination D

A.2 Chapter 3

f_{syb}	Sybil frequency
t_{mon}	Monitoring interval of the intrusion detection system in seconds
$n_{X,rec}$	Number of packets a node X received during t_{mon}
$n_{X,forw}$	Number of packets a node X forwarded during t_{mon}
w_{bal}	Factor to balance $n_{X,rec}$ and $n_{X,forw}$
$R_{X,Y}$	Rating for node Y determined by the intrusion detection system
$thres_{black}$	Threshold of R_Y for classification as black hole
r_{quar}	Radius of quarantined areas
t_{detect}	Time needed to detect a black hole
t_{reset}	Time after which quarantined areas are revoked
r_{trans}	Transmission range of nodes
A_{net}	Size of the network
l	Side length of the network area
d_{hop}	Average distance per hop
n_{total}	Total number of nodes in the network
n_{black}	Number of black hole nodes in the network
n_s	Number of nodes reached in step s of the ring search
ρ	Network density in nodes per area

A.3 Chapter 4

d_{center}	Distance to center of closest quarantined area

A.4 Chapter 5

$t_{retrans}$	Time between two consecutive retransmission attempts
$n_{retrans}$	Maximum number of retransmission attempts per packet

A.5 Chapter 6

K_i	i-th keyshare of Key K
$q(x)$	Polynomial used to generate keyshares
P_i	i-th peer
S_i	Partial signature generated with i-th keyshare
$S(m)$	Signature of Message m
L_i	i-th Lagrange factor

A.6 Chapter 7

n_{thres}	Number of partially signed certificates required to compute a valid certificate
p_{rep}	Probability with which a single peer answers a request
n_{keys}	Number of peers holding keyshares
n_{req}	Number of peers to which a request is sent
n_{rep}	Number of replies received to one request
$p(n_{rep})$	Probability for receiving exactly n_{rep} replies to one request
p_{succ}	Probability for receiving a sufficient (subject to n_{thres}) number of replies to one request

B Author's Publications

B.1 Publications as First Author

1. André König, Matthias Hollick, and Ralf Steinmetz. A Stochastic Analysis of Secure Joint Decision Processes in Peer-to-Peer Systems. In *Proceedings of the IEEE International Conference on Communications (ICC '09)*. 2009.

2. André König, Matthias Hollick, and Ralf Steinmetz. An Evaluation of Cooperative Decisions in Peer-to-Peer Systems - Mathematics vs. Testbed Studies. In *Proceedings of the 9th Würzburg Workshop on IP: Joint ITG and Euro-NF Workshop 'Visions of Future Generation Networks' (EuroView '09)*. 2009.

3. André König, Matthias Hollick, and Ralf Steinmetz. On the Implications of Adaptive Transmission Power for Assisting MANET Security. In *Proceedings of the Third IEEE International Workshop on Wireless Mesh and Ad Hoc Networks (WiMAN 2009)*. 2009.

4. André König, Aleksandra Kovacevic, and Ralf Steinmetz. Issues, Challenges, and Opportunities of Setting up Experimental Peer-to-Peer Facilities - A Personal Story. In *Proceedings of Panlab Workshop on Setup and Operation of Open Testbed Infrastructures in the Context of NGN and Future Internet - Status Quo and Quo Vadis in conjunction with 16. ITG/GI Fachtagung Kommunikation in Verteilten Systemen (KiVS '09)*. 2009.

5. André König, Daniel Seither, Matthias Hollick, and Ralf Steinmetz. An Analytical Model of Routing, Misbehavior, and Countermeasures in Mobile Ad Hoc Networks. In *Proceedings of the IEEE Global Communications Conference (GLOBECOM '09)*. 2009.

6. André König, Christian Gottron, Matthias Hollick, and Ralf Steinmetz. Harnessing Delay Tolerance to Increase Delivery Ratios in Mobile Ad Hoc Networks with Misbehaving Nodes. In *Proceedings of the Fourth IEEE International Workshop on Wireless and Sensor Networks Security (WSNS '08)*. 2008.

7. André König, Matthias Hollick, Tronje Krop, and Ralf Steinmetz. GeoSec: quarantine zones for mobile ad hoc networks. *Security and Communication Networks (Wiley SCN)*, 2:271–288, 2008.

8. André König, Matthias Hollick, and Ralf Steinmetz. Kooperative Sicherheitsmechanismen für Peer-to-Peer Systeme - Varianten und Erfolgsmodelle. In *Proceedings of 3. Essener Workshop 'Neue Herausforderungen in der Netzsicherheit' (EWNS '08)*. 2008.

9. André König, Matthias Hollick, and Ralf Steinmetz. Security for Future Wireless and Decentralized Communication Networks - Harnessing Cooperation, Space, and Time. In *Proceedings of the 8th Würzburg Workshop on IP: Joint EuroNF, ITC, and ITG Workshop 'Visions of Future Generation Networks' (EuroView '08)*. 2008.

10. André König, Ralf Ackermann, Matthias Hollick, and Ralf Steinmetz. Geographically Secure Routing for Mobile Ad Hoc Networks: A Cross-layer Based Approach. In *Proceedings of the Workshop on Long-Term and Dynamical Aspects of Information Security in conjunction with Emerging Trends in Information and Communication Security (ETRICS '06)*. 2006.

11. André König, Matthias Hollick, Johannes Schmitt, and Ralf Steinmetz. Sicherheit und Verfügbarkeit in mobilen Ad hoc Netzen - Ein geographischer, schichtenübergreifender Ansatz. In *Proceedings of 2. Essener Workshop 'Neue Herausforderungen in der Netzsicherheit' (EWNS '06)*. 2006.

B.2 Publications as Coauthor

12. Christian Gottron, André König, and Ralf Steinmetz. A Cross-layer Approach Towards Robustness of Mobile Peer-to-Peer Networks. *Submitted to the 45th IEEE International Conference on Communications (ICC '11)*, 2011.

13. Christian Gottron, André König, and Ralf Steinmetz. A Cross-Layer Approach for Increasing Robustness of Mobile Peer-to-Peer Networks. In *Proceedings of the 'Security in NGNs and the Future Internet' Workshop collocated with the 3rd Future Internet Symposium 2010 (FIS '10)*. 2010.

14. Christian Gottron, André König, and Ralf Steinmetz. A Survey on Security in Mobile Peer-to-Peer Architectures - Overlay-based vs. Underlay-based Approaches. *Accepted for MDPI Future Internet special issue 'Network vs. Application Based Solutions for NGN'*, 2010.

15. Christian Gottron, André König, and Ralf Steinmetz. A Testbed-based Analysis of the Incorrect Lookup Routing Attack on the Pastry DHT. In *Proceedings of the 10th Würzburg Workshop on IP: Joint ITG, ITC, and Euro-NF Workshop 'Visions of Future Generation Networks' (EuroView '10)*. 2010.

16. Christian Gottron, Pedro Larbig, André König, Matthias Hollick, and Ralf Steinmetz. Testbed Evaluation eines Black Hole-Angriffes auf ein Ad hoc Netz. In *Proceedings of 4. Essener Workshop 'Neue Herausforderungen in der Netzsicherheit (EWNS '10)*. 2010.

17. Christian Gottron, Pedro Larbig, André König, Matthias Hollick, and Ralf Steinmetz. The Rise and Fall of the AODV Protocol: A Testbed Study on Practical Routing Attacks. In *Proceedings of the 35th IEEE Conference on Local Computer Networks (LCN '10) (accepted for presentation)*. 2010.

18. André Miede, Nedislav Nedyalkov, Christian Gottron, André König, Nicolas Repp, and Ralf Steinmetz. A Generic Metamodel for IT Security - Attack Modeling for Distributed Systems. In *Proceedings of the Fifth International Conference on Availability, Reliability, and Security (ARES '10)*. 2010.

19. Ralf Steinmetz and André König. On Simulations, Models, and Testbeds - Possibilities and Pitfalls. *Submitted to Elsevier Performance Evaluation Special Issue on Selected Papers of MSWiM 2009*, 2010.

20. Christian Gottron, André König, Matthias Hollick, Sonja Bergsträßer, Thomas Hildebrandt, and Ralf Steinmetz. Quality of Experience of Voice Communication in Large-Scale Mobile Ad Hoc Networks. In *Proceedings of the Second IFIP Wireless Days 2009*. 2009.

21. André Miede, Christian Gottron, André König, Nedislav Nedyalkov, Nicolas Repp, and Ralf Steinmetz. Cross-organizational Security in Distributed Systems. Technical report, Multimedia Communications Lab (KOM), TU Darmstadt, 2009.

22. Johannes Schmitt, Oliver Heckmann, André König, Matthias Hollick, and Ralf Steinmetz. ENUM-Erweiterung zur Sicherung der Kommunikation zwischen VoIP-Infrastrukturen. In *Proceedings of 2. Essener Worshop 'Neue Herausforderungen in der Netzsicherheit' (EWNS '06)*. 2006.

C Curriculum Vitae

Personal Information

André König
Tannenweg 9
64560 Riedstadt
Germany

Education

10/1998–09/2005	Diploma in Computer Science (with honors), Technische Universität Darmstadt, Darmstadt, Germany
08/1997–09/1998	Alternative service, German Red Cross, Darmstadt, Germany
08/1988–07/1997	High school 'Gymnasium Gernsheim', Gernsheim, Germany
08/1984–07/1988	Elementary school, Biebesheim, Germany

Professional Experience

08/1998–09/2005	Consultant for network and system design, Drumm Systemhaus, Gernsheim, Germany

Academic Experience

since 10/2005	Doctoral candidate and research staff member, Multimedia Communications Lab (KOM), Technische Universität Darmstadt
since 10/2005	Teaching assistant for the masters level course 'Communication Networks I', Multimedia Communications Lab (KOM), Technische Universität Darmstadt
since 10/2005	Tutor for the seminar 'Advanced Topics of Future Internet Research', Multimedia Communications Lab (KOM), Technische Universität Darmstadt
since 04/2006	Tutor for the seminar 'Security in Mobile Ad Hoc Networks', Multimedia Communications Lab (KOM), Technische Universität Darmstadt, in cooperation with the Fraunhofer Institute for Computer Graphics Research
since 10/2008	Local project supervisor of the research project 'German Lab (G-Lab), National Platform for Future Internet Studies' funded by the German Ministry of Education and Research, Multimedia Communications Lab (KOM), Technische Universität Darmstadt
since 04/2009	Head of the research group 'Network Security', Multimedia Communications Lab (KOM), Technische Universität Darmstadt
09/2009–12/2009	Research scholar, Multimedia Operating System and Networking Group lead by Professor Klara Nahrstedt, Department of Computer Science, University of Illinois at Urbana-Champaign (UIUC), funded by the German Academic Exchange Service (DAAD) (09/2009–11/2009)
01/2006–10/2007	Local project supervisor of the research project 'SicAri - A Security Architecture and its Tools for Ubiquitous Internet Usage' funded by the German Ministry of Education and Research, Multimedia Communications Lab (KOM), Technische Universität Darmstadt

D Erklärung laut §9 der Promotionsordnung

Ich versichere hiermit, dass ich die vorliegende Dissertation allein und nur unter Verwendung der angegebenen Literatur verfasst habe. Die Arbeit hat bisher noch nicht zu Prüfungszwecken gedient.

Darmstadt, 2011

Dipl.-Inform. André König